An Analysis of the *daf-2* insulin/ IGF-1 receptor gene in *Caenorhabditis elegans*

Dhaval Subhas Patel

University College London

PhD Thesis

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

The *daf-2* gene is a key regulator of growth, metabolism and longevity in the nematode *Caenorhabditis elegans*. The DAF-2 receptor functions in a pathway that is analogous to mammalian insulin/ insulin-like growth factor signalling and determines whether animals proceed with full reproductive development or arrest in a long-lived diapausal state known as the dauer larva. Temperature-sensitive hypomorphic mutants of *daf-2* constitutively arrest as dauers when raised at non-permissive temperatures. At permissive temperatures the animals develop into adults that are long-lived compared to wild-type adults. These alleles of *daf-2* can be separated into two distinct classes (1 and 2) based on their pleiotropic phenotypes. In addition to their phenotypic differences, the two classes also differ in their epistatic interactions with other genes involved in dauer formation, suggesting that the DAF-2 receptor has multiple signalling outputs.

In this thesis I have investigated the nature of the *daf-2* allele class difference using a range of methods, including sequence analysis and homology modelling of mutant receptors. This generated the prediction that signalling flux through the receptor is a determinant of the class difference, with class 2 alleles having an asymmetrical alteration in signal transduction through DAF-2. Experimental testing of these predictions suggest that some phenotypes such Eat and Unc may be associated with asymmetrical signalling, while others such as early larval arrest may be correlated with a reduction in receptor level at the plasma membrane.

A comparative analysis of the DAF-2 receptor with its homologues from *Caenorhabditis briggsae*, *Caenorhabditis remanei* and the parasitic nematode *Brugia malayi* suggests the large C-terminal extension in the *Caenorhabditis* species, which shorter in *Brugia* and not present in vertebrates, is an adaptive trait that has evolved by exon duplication for rapid growth and development and may contribute to the shortevity of these species. In addition, I have also performed an analysis of *ins-7* and *ins-35*, two putative ligands of DAF-2 that are differentially regulated by the transcription factor DAF-16, using RNAi. Knockdown of both these genes lead to slight lifespan extension (~10-20%) at 20°C in all genetic backgrounds and slight suppression (~10%) at 25°C in long-lived *daf-2* mutant backgrounds compared to controls. This suggests that INS peptides may function as agonists a 20°C and antagonists at 25°C, and that this behaviour may be independent of their transcriptional regulation.

Declaration

I declare that the work presented in this thesis is my own except where duly noted

Dhaval S. Patel

Dedicated to

My parents Subhas and Kalpana

My sisters Kruti and Smruti

And All my Friends and Family

Acknowledgements

I would like to thank the following people:

Acely Garza-Garcia, Paul Driscoll, Nachiket Nadkarni, Sindhuja Kadambi, Edward Campbell, Peter Ottensmeyer, Manoj Nanji, Joshua McElwee, Alice Smith, Mari-Wyn Burley, Cynthia Kenyon, James H. Thomas, Rick Maizels, Yvonne Harcus, Murray Selkirk, David Guiliano and Mark Blaxter for their direct contributions to this project.

Diana McCulloch, Glenda Walker, Dave Weinkove, Adelaide Raimundo and Michelle Keaney: Past and Present members of the Gems Lab.

All members of the Partridge and Chapman Labs. Especially Giovanna Vinti for keeping the labs running and Susan Broughton for letting me DJ at her 40th birthday party.

Will Mair, Stuart Wigby, Jon Linklater, Melanie West, Martin Sikora and Fede Calboli for providing many amusing moments during my PhD.

I would like to thank the BBSRC for funding this project.

I would like to thank my second supervisor Sally Leevers.

Finally, I would like to thank my supervisor David Gems for his patience, support and good humour throughout my PhD.

Table of Contents

Table of Tables

Table of Figures

Chapter 1

Introduction

1.1 Insulin/ insulin-like growth factor signalling (IIS) and ageing

This thesis describes an analysis of the *daf-2* gene, and the protein it encodes, in the nematode *Caenorhabditis elegans*. The DAF-2 receptor is a member of the insulin receptor (IR) superfamily of proteins. Even though Kimura *et al* elucidated the sequence of the *daf-2* gene in 1997, a number of questions concerning its confusing genetics and the aspects of its predicted amino acid sequence remain. The DAF-2 receptor sits atop the insulin/insulin-like growth factor signalling (IIS) pathway in *C. elegans,* in which there is currently intense interest, as this pathway is a major regulator of longevity. Hypomorphic mutations in several components of the IIS pathway in *C. elegans*, including *daf-2*, lead to lifespan extension (Figure 1.1 and 1.7) (Friedman and Johnson, 1988b).

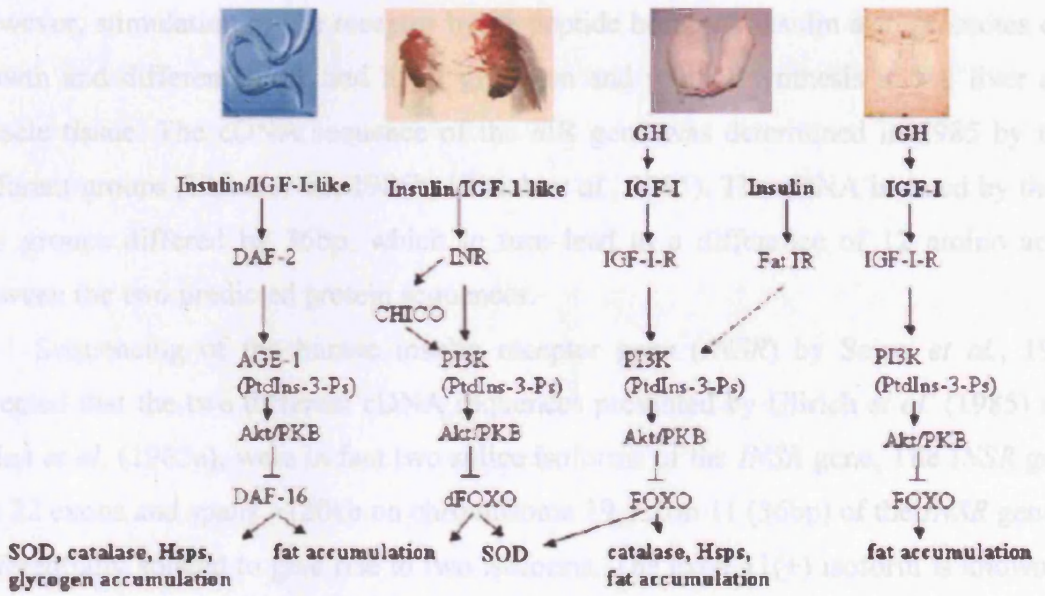
C. elegans is not the only species in which the IIS pathway is a regulator of longevity. Mutations in the *Drosophila melanogaster* IIS pathway can also cause extension of lifespan. Like DAF-2, mutations in dInR (*D. melanogaster* insulin receptor) can extend lifespan (Figure 1.1). However, unlike DAF-2, these mutations only lead to lifespan extension in heteroallelic combinations, as most alleles of *dinr* are homozygous lethal (Tatar *et al.*, 2001). A second component of the IIS pathway that leads to lifespan extension in *D. melanogaster* is CHICO, which is the fly homologue of the mammalian insulin receptor substrate (IRS) proteins (Clancy *et al.*, 2001). Interestingly, the mutation in CHICO is semi-dominant as heterozygotes are also long-lived.

The IIS pathway has recently been found to play a role in regulation of longevity in mice. However, analysis of the IIS pathway in mice is complicated by the fact that the mouse genome, like that of all mammals, encodes three different members of the IR superfamily of proteins. These are the insulin receptor, insulin-like growth factor-1 receptor (IGF-1R) and the insulin-related receptor (IRR). Female mice that are heterozygous for an *Igf1R* gene knockout live 33% longer than wild-type controls (Holzenberger *et al.*, 2003). The insulin receptor also seems to play a role in regulation of longevity in mice. The fat-specific insulin receptor knockout (FIRKO) mouse is long-lived, with an 18% increase in mean lifespan (Bluher *et al.*, 2003), which suggests that the IGF-1R has a greater role in lifespan regulation in mice than IR. The Prop-1 (Ames) and Pit-1 (Snell) dwarf mice are mutants of two transcription factors that control

pituitary development (reviewed (Katic and Kahn, 2005)). These mice are unable to release growth hormone (GH), which is needed to simulate IGF-1 release by the liver. This leads to a reduction in circulating IGF-1 levels, which in turn causes lifespan extension (see Figure 1.1).

Potentially the regulation of longevity by IIS extends up the evolutionary tree to humans. Several groups have shown that certain polymorphic alleles of genes in the IIS pathway are enriched in populations of long-lived individuals. One such allele is the *IGF-IR* A^+ of the *h*IGF-1R (Bonafe *et al.*, 2003). This variant seems to reduce circulating levels of IGF-1 in the individuals who carry it. Recently, an allele of GH that seems to reduce height in carriers and decrease their mortality has been identified (van Heemst *et al.*, 2005).

Figure 1.1 Conservation of IIS regulated longevity pathway



Schematic representation of the conserved elements of the IIS pathway in *C. elegans*, *D. melanogaster*, *M. musculus* and humans. Insulin/ IGF-1-like peptides in all four species bind to and activate insulin-related receptors (DAF-2 in *C. elegans*, dInR in *D. melanogaster*, IGF-1R and IR in *M. musculus* and humans), which leads to the activation of a conserved PI 3-kinase pathway. This pathway negatively regulates transcription factors from the FOXO family (DAF-16 in *C. elegans*, dFOXO in *D. melanogaster*, FOXO in *M. musculus* and humans). When FOXO transcription factors are active they up-regulate anti-oxidant proteins, such as SOD and catalase, as well as fat accumulation.

1.2 The human insulin receptor

The human insulin receptor (*h*IR) is the best-characterised member of the IR superfamily of proteins, of which DAF-2 is a member. Below I discuss the structure of the *INSR* gene and the predicted structure of *h*IR protein. I also describe the clinical syndromes that arise from mutations in *h*IR, as well as classification system that is used to describe the molecular effects of the mutation on the receptor (Figure 1.4).

1.2.1 The human insulin receptor gene

The human insulin receptor is transmembrane receptor tyrosine kinase (RTK) that is a key regulator of growth and metabolism (Le Roith and Zick, 2001; Saltiel and Kahn, 2001). The principal function of the receptor is to regulate blood glucose concentration. However, stimulation of the receptor by its peptide hormone insulin also promotes cell growth and differentiation, and lipid, glycogen and protein synthesis in fat, liver and muscle tissue. The cDNA sequence of the *h*IR gene was determined in 1985 by two different groups (Ebina *et al.*, 1985b; Ullrich *et al.*, 1985). The cDNA isolated by these two groups differed by 36bp, which in turn lead to a difference of 12 amino acids between the two predicted protein sequences.

Sequencing of the human insulin receptor gene (*INSR*) by Seino *et al.*, 1989 revealed that the two different cDNA sequences presented by Ullrich *et al.* (1985) and Ebina *et al.* (1985a), were in fact two splice isoforms of the *INSR* gene. The *INSR* gene has 22 exons and spans >120kb on chromosome 19. Exon 11 (36bp) of the *INSR* gene is differentially spliced to give rise to two isoforms. The exon 11(+) isoform is known as the A isoforms and represents the sequence identified by Ebina *et al.* (1985a). The exon 11(-) isoform is known as the B isoform and represents the sequence the presented by Ullrich *et al.* (1985). Throughout this thesis the numbering of *h*IR residues will use the same format as Ebina *et al.* (1985a).

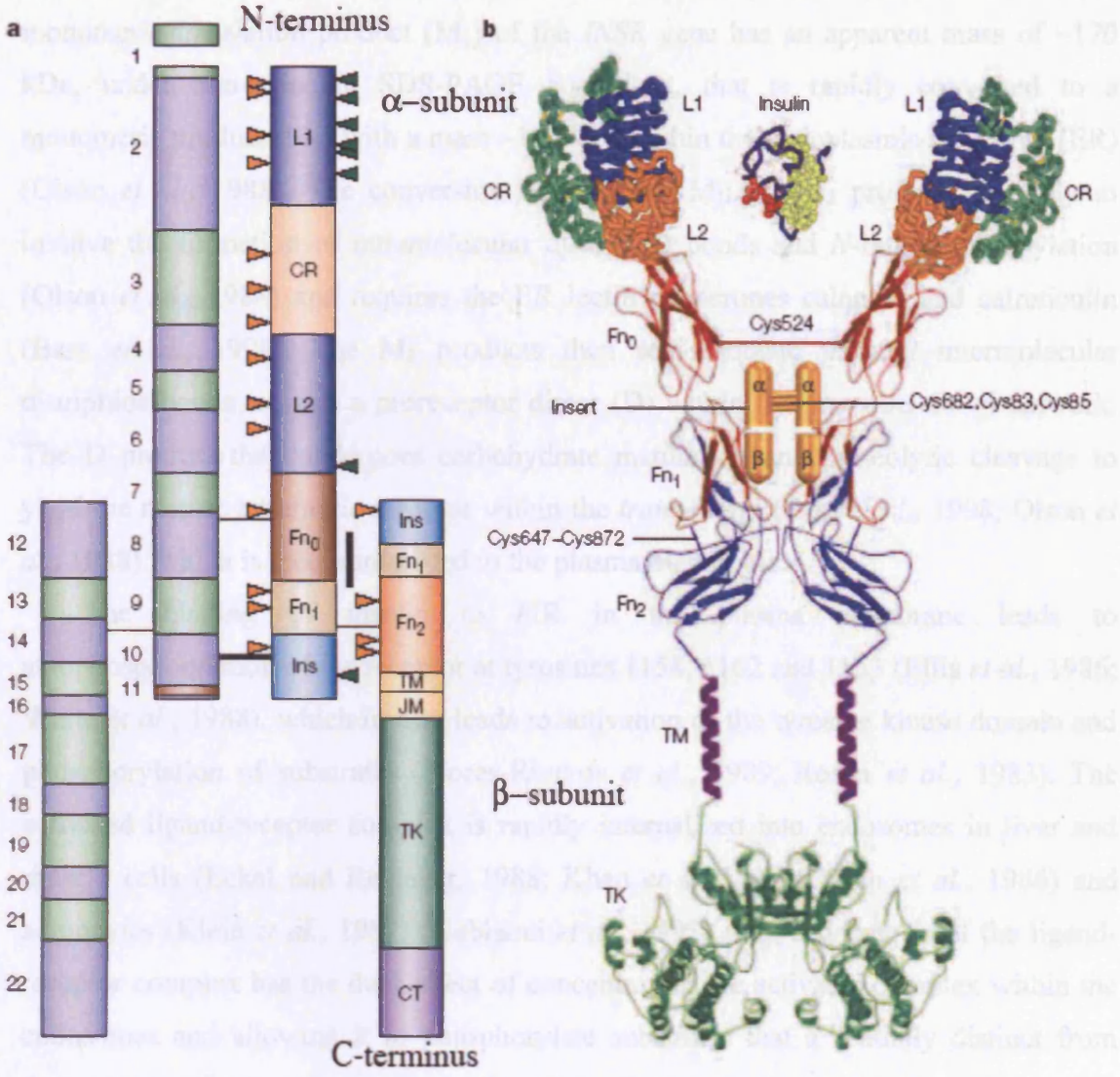
1.2.2 Structure and processing of the *h*IR protein

The mature *h*IR protein is a heterotetramer composed of two covalently linked chains. Each chain is itself a dimer that is composed of an α-subunit and β-subunit, these subunits are derived from a single proreceptor and are linked by a disulphide bond between cysteines 647 and 872 (Cheatham and Kahn, 1992; Lu and Guidotti, 1996; Sparrow *et al.*, 1997). The α-β chains are linked by four disulphide bonds between the equivalent cysteine residues 524, 682, 683 and 685 in adjacent α-subunits (Schaffer and Ljungqvist, 1992; Sparrow *et al.*, 1997; Wu and Guidotti, 2004). The final subunit structure of the mature *h*IR protein is $\alpha_2\beta_2$ (Figure 1.2).

The α-subunit of each receptor chain is entirely extracellular, while the β-subunit contains a portion that is extracellular, the transmembrane domain and the cytoplasmic tyrosine kinase domain. The α-subunit contains two repeats of a leucine rich domain (L1 and L2) (Bajaj *et al.*, 1987; Ward and Garrett, 2001) separated by a cysteine rich (CR) domain similar to that found in the epidermal growth factor receptor (Bajaj *et al.*, 1987; Ward *et al.*, 1995). This is followed by the Fn0 domain (Marino-Buslje *et al.*, 1998; Mulhern *et al.*, 1998; Ward, 1999), which is a fibronectin type III domain. The α-subunit also contains the first half of the Fn1 domain, which is also a fibronectin type III domain; the other half of this domain is part of the β-subunit. The Fn1 domain is interrupted by an insert region that spans the α-subunit and β-subunits that contains a tetrabasic endoproteolytic cleavage site.

The β-subunit contains the other half of the Fn1 domain followed by a third fibronectin type III domain, known as Fn2 (Schaefer *et al.*, 1992). The transmembrane domain then follows these domains and the rest of the β-subunit, which contains the tyrosine kinase (TK) domain and a short C-terminal (CT) extension, is cytoplasmic (Ebina *et al.*, 1985b; Ullrich *et al.*, 1985).

Figure 1.2 Structure of insulin receptor



Panel **a** shows a cartoon of the $\alpha_2\beta_2$ structure of the insulin receptor. The left half of the diagram shows the boundaries of the 22 exons of the insulin-receptor gene. The right half of the diagram outlines the predicted boundaries of the protein modules. Orange arrowheads indicate *N*-glycosylation sites. Green arrowheads indicate ligand-binding 'hot spots' that have been identified by single-amino-acid site-directed mutagenesis. Panel **b** shows the domain organization of the insulin receptor. The diagram depicts a model of predicted or actual modular structures. The three-dimensional structure of the insulin molecule is also shown. Image taken from (De Meyts and Whittaker, 2002).

The A and B splice isoforms of the *INSR* gene give rise to *h*IR proreceptors that are 1382 and 1370 amino acids long, respectively. Each proreceptor undergoes series of processing steps that leads to the formation of the mature *h*IR protein. The initial monomeric translation product (M_1) of the *INSR* gene has an apparent mass of ~170 kDa, under non-reducing SDS-PAGE conditions, that is rapidly converted to a monomeric product (M_2) with a mass ~190 kDa within the endoplasmic reticulum (ER) (Olson *et al.*, 1988). The conversion between the M_1 and M_2 products is likely to involve the formation of intramolecular disulphide bonds and *N*-linked glycosylation (Olson *et al.*, 1988) and requires the ER lectin chaperones calnexin and calreticulin (Bass *et al.*, 1998). The M_2 products then self-associate through intermolecular disulphide bonds to form a proreceptor dimer (D) within the pre-*trans*-Golgi network. The D product then undergoes carbohydrate maturation and proteolytic cleavage to yield the mature tetrameric receptor within the *trans*-Golgi (Bass *et al.*, 1998; Olson *et al.*, 1988), which is then transported to the plasma membrane.

The binding of insulin to *h*IR in the plasma membrane leads to autophosphorylation of the receptor at tyrosines 1158, 1162 and 1163 (Ellis *et al.*, 1986; White *et al.*, 1988), which in turn leads to activation of the tyrosine kinase domain and phosphorylation of substrates (Flores-Riveros *et al.*, 1989; Rosen *et al.*, 1983). The activated ligand-receptor complex is rapidly internalized into endosomes in liver and muscle cells (Eckel and Reinauer, 1988; Khan *et al.*, 1989; Khan *et al.*, 1986) and adipocytes (Klein *et al.*, 1987) (Kublaoui *et al.*, 1995). The endocytosis of the ligand-receptor complex has the dual effect of concentrating the activated complex within the endosomes and allowing it to phosphorylate substrates that a spatially distinct from those at the plasma membrane (reviewed in Bevan *et al.*, 1996). The endosome is acidified by active transport of protons into the endosomal lumen through an ATP-driven proton pump, which causes the insulin ligand to dissociate from its receptor and be degraded by an insulin-specific protease (reviewed in Knutson, 1991 and Bevan *et al.*, 1996). The unbound receptor is then dephosphorylated by a phosphotyrosine phosphatase, which might be associated with the endosomal membrane (Faure *et al.*, 1992), before either being recycled to the plasma membrane via the Golgi apparatus or degraded via a lysosomal pathway (reviewed in Knutson, 1991).

1.2.3 Signalling through the insulin receptor

The insulin receptor is a key regulator of metabolism and growth and one of its major signalling outputs is the PI 3-kinase complex (Shepherd *et al.*, 1995). The insulin receptor does not directly interact with the PI 3-kinase complex; instead activation of the tyrosine kinase domain of the receptor leads to phosphorylation of the insulin receptor substrate proteins (IRS-1 to -4), which then interact with p85 (the regulatory subunit of the PI 3-kinase complex) (Sun *et al.*, 1991; White *et al.*, 1985). Activation of the PI 3-kinase complex results in the phosphorylation of the lipid phosphatidylinositol-(4,5)-bisphosphate (PIP2) to phosphatidylinositol-3,4,5-triphosphate (PIP3) (Whitman *et al.*, 1988). Elevated levels of PIP3 act as a secondary messenger that activates multiple signalling pathways leading to an increase in glycogen, lipid and protein synthesis, as well as cell growth and differentiation (see Figure 1.3) (reviewed in Saltiel and Khan, 2001).

In addition to signalling through the PI 3-kinase complex, the insulin receptor also activates the Ras/ Mitogen-activated protein (MAP) kinase pathway and the CAP/ Cbl pathway (Pirola *et al.*, 2004; Saltiel and Kahn, 2001). The CAP/ Cbl pathway is also involved in the regulation of glucose uptake, while the Ras/ MAP kinase pathway is involved with cell proliferation and differentiation (see Figure 1.3). Activation of Ras signalling by the insulin receptor involves recruitment of the Grb2.Sos complex, which can bind to either phosphorylated IRS-1 or Shc. Activation of Ras by both the insulin receptor and the IGF-1 receptor is predominately transduced via Shc and requires receptor internalization (Ceresa *et al.*, 1998; Chow *et al.*, 1998; Sasaoka *et al.*, 1994).

Figure 1.3 Signalling through the insulin receptor

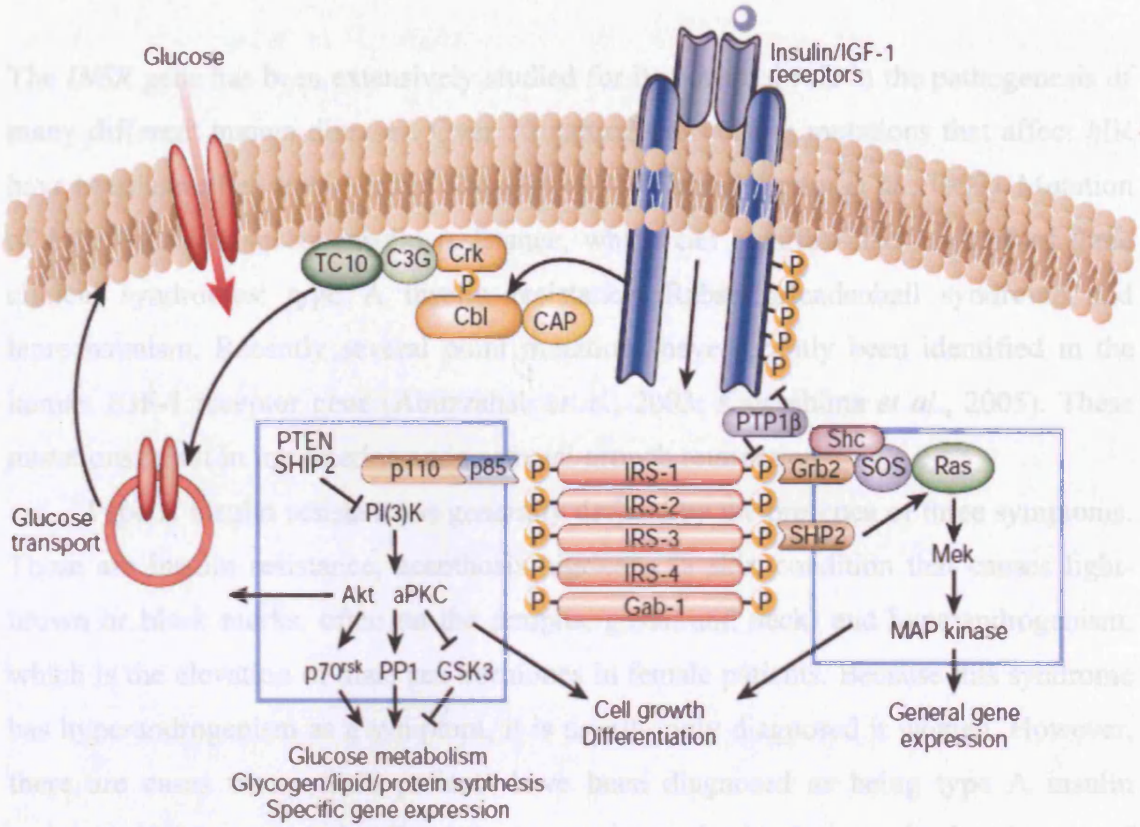


Diagram showing the three main pathways activated by both the insulin and IGF-1 receptors. Both of these receptors mediate their signal transduction through phosphorylation of adaptor proteins, such IRS-1 to -4, Gab-1, CAP and Shc, which in turn recruit different effector proteins. The major output of both the insulin and IGF-1 receptor is the PI 3-kinase complex. Both receptors also can activate the Ras/ MAP kinase pathway, predominately through Shc, which requires receptor internalization. The CAP/ Cbl pathway is involved in regulating glucose uptake and functions additively to the PI 3-Kinase pathway. Image taken from (Saltiel and Kahn, 2001).

1.2.4 Mutations affecting the human insulin receptor

The *INSR* gene has been extensively studied for its potential role in the pathogenesis of many different human diseases. Over 50 naturally occurring mutations that affect *h*IR have been identified (reviewed in Krook *et al.*, 1997a and Taylor *et al.*, 1992). Mutation of *h*IR results in severe insulin resistance, which can manifest itself as one of three clinical syndromes: type A insulin resistance, Rabson-Mendenhall syndrome and leprechaunism. Recently several point mutations have recently been identified in the human IGF-1 receptor gene (Abuzzahab *et al.*, 2003; Kawashima *et al.*, 2005). These mutations result in intrauterine and postnatal growth retardation.

Type A insulin resistance is generally defined by the presence of three symptoms. These are insulin resistance, acanthosis nigricans (a skin condition that causes light-brown or black marks, often on the armpits, groin, and neck) and hyperandrogenism, which is the elevation of male sex hormones in female patients. Because this syndrome has hyperandrogenism as a symptom, it is usually only diagnosed it women. However, there are cases where male patients have been diagnosed as being type A insulin resistant, if they present insulin resistance and acanthosis nigricans in the absence of other clinical symptoms (Taylor *et al.*, 1992).

Rabson-Mendenhall syndrome is defined by the presence of several different symptoms. Firstly, patients have extreme insulin resistance. They also have acanthosis nigricans, as well as abnormalities of their teeth and nails, and pineal hyperplasia. Rabson-Mendenhall syndrome is usually regarded as being intermediate to type A insulin resistance and leprechaunism. Leprechaunism and Rabson-Mendenhall syndrome have very similar clinical symptoms, making it difficult to separate the two. The main difference between diagnoses of either syndrome is the age at which the patient dies. Patients with Rabson-Mendenhall syndrome are not expected to reach puberty, while patients with leprechaunism usually die within the first year of infancy.

Mutations that affect the *INSR* gene can be classified by their effect on the receptor (Taylor *et al.*, 1992). There are five different points at which mutations might affect the receptor; these are listed in Table 1.2.1. Type 1 mutations affect transcription of the *INSR* mRNA. Type 2 mutations affect transport of the receptor to plasma membrane, through defective processing of the proreceptor. Type 3 mutations affect the insulin binding affinity of the receptor. Type 4 mutations affect the signalling output of

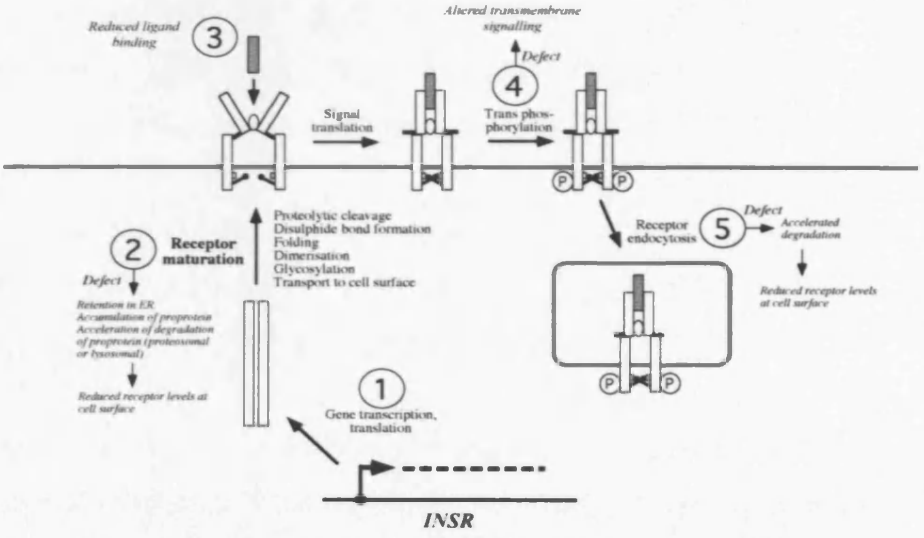
the receptor, usually by altering kinase activity. Type 5 mutations affect the number of receptors at the cell surface by interfering with receptor endocytosis and its subsequent recycling or degradation. Mutations can actually result in more than one type of defect in the receptor. For example, the R252C mutation of *h*IR reduces the transport of the receptor to cell surface (type 2), reduces insulin binding (type 3), alters the signalling output of the receptor (type 4) and reduces the internalization of the receptor (type 5) (Hamer *et al.*, 2002).

Table 1.2.1 Classification of mutations in *INSR* gene/ *h*IR protein

Type of mutation	Synthesis and translation of mRNA	Transport to plasma membrane	Insulin-binding affinity	Transmembrane signalling/ kinase activity	Endocytosis, recycling and degradation of receptor
1	X				
2		X			
3			X		
4				X	
5					X

Classification system for mutations affecting *h*IR. The crosses mark the biological time point at which each mutation affects receptor function. Type 1 mutations affect biosynthesis of the receptor, type 2 affect maturation and transport to the cell surface, type 3 affect insulin-binding affinity, type 4 affect signal transduction and type 5 affect receptor turnover.

Figure 1.4 Types of defect affecting *INSR* gene/ *h*IR protein



A diagrammatic representation of Table 1.2.1

1.2.5 Biochemical analysis of mutations in the insulin receptor

The sequencing of the *INSR* cDNA by Ebina *et al.* (1985a) and Ullrich *et al.* (1985) enabled the design of expression constructs for both wild-type and mutant receptors that could be expressed in a heterologous cell system. Typical cell lines used in these heterologous expression systems are Chinese hamster ovary (CHO), COS 7 (simian kidney cells), NIH 3T3-L1 adipocytes (murine fibroblast-derived cell line), Rat-1 fibroblasts or HTC hepatoma cells (reviewed in Tavare and Siddle, 1993). CHO cells were the first cell line to be used for heterologous expression of wild-type *h*IR (Ebina *et al.*, 1985a). These cells have a low level of expression of endogenous insulin receptors, typically ~2800 per cell (Ebina *et al.*, 1985a), and have well know insulin-binding kinetics (Podskalny *et al.*, 1984). Similarly, all the cell lines also have low expression levels of endogenous receptors, ~2000-5000 per cell, there are no receptor-null cell lines so endogenous receptor effects must be considered when interpreting data from these cell lines (Tavare and Siddle, 1993).

1.3 Introduction to *Caenorhabditis elegans*

In this section I describe the features of *C. elegans* that make it a useful organism for the study of ageing. I then present the IIS pathway in worms and the interesting genetics of *daf-2*.

1.3.1 *Caenorhabditis elegans* as a model organism

Caenorhabditis elegans is a convenient model organism for the study of a broad range of avenues of biological investigation, particularly ageing research. The reproductive cycle of *C. elegans* is three days from egg to fertile adult (depending on the temperature at which animals are cultured), and the adult lifespan is approximately two weeks. Adults are only ~1.2mm in length and so can be grown and handled in large quantities. They are also bacteriovores and can be maintained on many different types of bacteria, but are usually grown on *E. coli* strain OP50 (Brenner, 1974), either on agar plates or in liquid culture. One the most convenient aspects of *C. elegans* biology is its mode of sexual reproduction. The majority of animals tend to be self-fertilising hermaphrodites, which allows the creation of isogenic lines, and facilitates the formation of homozygous mutants. However, males are produced at low frequency, which are able to mate with the hermaphrodites. This enables backcrossing and the ability to make strains carrying multiple mutations. The entire cell lineage of *C. elegans* has been determined (Sulston and Horvitz, 1977), which makes this organism very attractive for developmental studies. Finally, strains can be frozen and stored for many years at -70°C, which allows the building of large strain collections without having to continuously culture the animals.

The genetics and genomics of *C. elegans* are also very well developed. The entire genome has been sequenced (Consortium, 1998), and is currently the only metazoan genome to be truly finished (i.e. no gaps) (Hodgkin, 2005). The genome contains over 22,000 predicted genes, and is available in a public database curated by the Wormbase consortium (www.wormbase.org). The ease of mutagenesis in *C. elegans*, with chemical mutagens such as ethyl methane sulfonate, has resulted in many different and well-characterised mutants that are freely available from the *Caenorhabditis* Genetics

stock Center (CGC), maintained at the University of Minnesota (http://biosci.umn.edu/CGC/CGChomepage.htm).

1.3.2 Genetics and nomenclature

Genetic loci in *C. elegans* are given names consisting of three or four italicized letters, a hyphen, and an italicized Arabic number, e.g., *daf-2* or *rskn-1*. This can be followed by an italicized Roman numeral to indicate the chromosome on which the gene is located, e.g., *daf-2 III* or *rskn-1 I*. The three or four letter abbreviations can refer either to the mutant phenotype of the gene (*da*uer *f*ormation abnormal, in the case of *daf-2*) or the product of the gene (*RSK*-p*N*inety, RSK-p90 kinase homolog, in the case of *rskn-1*). Alleles of genes are given in brackets after the gene name, with no spaces and in italics, as combination of the lab allele designation and an allele number. For example, the lab allele designation for the Gems lab is *wu*, so a *daf-2* mutant isolated in our lab would be *daf-2(wu123)*. Phenotypes are generally written as three letter abbreviations, these have a capital letter and are non-italicized, e.g. Unc (uncoordinated) or Egl (egg-laying defective). For a more thorough list of nomenclature guidelines see the CGC guidelines webpage (http://biosci.umn.edu/CGC/Nomenclature/nomenguid.htm).

1.3.3 Overview of *C. elegans* biology and its lifecycle

Caenorhabditis elegans is a free-living terrestrial nematode (Riddle *et al.*, 1997; Wood, 1988). There are two different sexes: hermaphrodites and males, which occur at low frequency in laboratory cultures. Sex determination within this species is based on the autosome to sex chromosome ratio; hermaphrodites are XX and males are X0. Adult hermaphrodites have 959 somatic nuclei and males have 1031.

C. elegans proceeds through four larval stages during its reproductive life cycle, from egg to sexually mature adult. These are stages are referred to as L1, L2, L3 and L4 (Figure 1.5); each stage is punctuated by a moult of the collagen cuticle surrounding the worm. Under optimal growth conditions the reproductive life cycle is complete in about 3 days; wild type worms then have a maximum life span of 2-3 weeks. However, in poor growth conditions such as overcrowding or high temperature the worms are able to

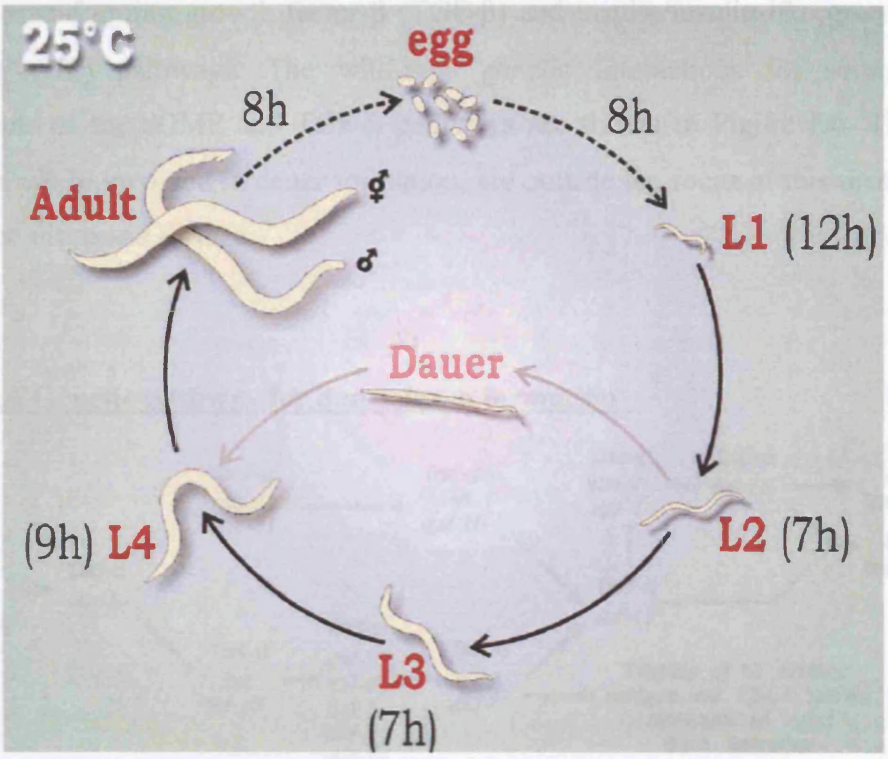
moult to an alternative third stage known as the dauer larva (Cassada and Russell, 1975).

The dauer larvae are arrested in a diapausal state that has several behavioural, morphological and metabolic adaptations that allows the worms to survive harsh environments and to resume normal growth once conditions improve. The decision to enter the dauer stage is made during the L1 stage and is based on a series of environmental cues (reviewed in Riddle and Albert, 1997). The first of these cues is concentration of a dauer pheromone in the environment, which is constitutively released by the animals, and so represents the population density of the worms. Recently, the structure dauer pheromone 'daumone' was chemically determined and found to be a heptanoic acid derivative (Jeong *et al.*, 2005). Food availability modulates the response to daumone, such that high daumone to low food ratios promote dauer formation, while the opposite conditions promote full reproductive development and recovery of dauer larva. Temperature also modulates the response to the daumone/food ratio, such that higher temperatures promote dauer formation at lower daumone/food ratios.

The worms enter the dauer larval stage via an extended L2 stage, known as L2d. The L2d stage usually last 13 hours, as opposed to 7 for the normal L2 stage. During the L2d stage the worm prepares to enter the dauer stage by accumulating fat in its intestinal and hypodermal cells. The L2d stage does not commit the worm to dauer larva formation, should growth conditions improve the L2d animals can moult to the normal L3 stage.

The worms are able to survive as dauer larvae for up to 70 days whilst retaining the ability to resume normal development and become fertile adults (Klass and Hirsh, 1976). The length of time spent as a dauer larva has no effect on the post-dauer maximum life span of the adult worm, for this reason the dauer larva is referred to as non-ageing (Klass and Hirsh, 1976). However, it is not yet known whether the dauer larva is truly non-ageing or whether the effects of ageing in the dauer larva are reversed before resuming normal growth.

Figure 1.5 The lifecycle of *C. elegans*



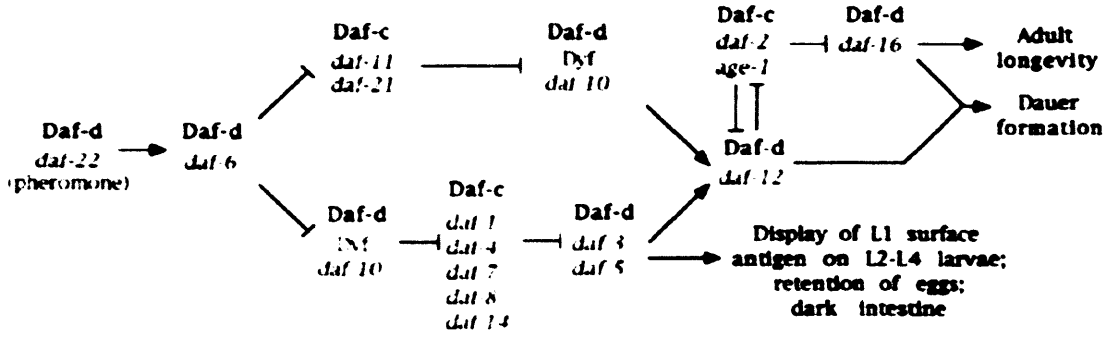
The lifecycle of *C. elegans* takes ~3 days from egg to fertile adult in favourable conditions at 25°C. In unfavourable conditions the worms can enter an alternate third stage known as the dauer larva. This image is taken from (Braeckman *et al.*, 2001).

1.3.4 The genetics of dauer formation

Over 30 genes have been identified as controlling dauer formation in *C. elegans* (Riddle and Albert, 1997), most of these are known as *daf* genes (*da*uer *f*ormation abnormal). The *daf* genes fall into two phenotypic groups, these are known as Daf-c and Daf-d. The Daf-c genes give rise to worms that constitutively form dauers even at low population density and with abundant food. Alleles of the Daf-c genes can either be non-conditional, in which case they form dauers independently of environmental cues, or temperature-sensitive (ts), in which case mutants only form dauers when shifted to restrictive temperatures. The Daf-d gene mutants are unable to enter the dauer larva stage even in restrictive environmental conditions such as over-crowding and starvation.

These genes have been assembled into 3 different but possibly interacting pathways using epistatic analysis. These are the cyclic guanosine monophosphate (cGMP), transforming growth factor-β (TGF-β) and insulin/insulin-like growth factor signalling (IIS) pathways. The wild-type genetic interactions for some of the components of the cGMP and TGF-β pathways are shown in Figure 1.6. These two pathways, while involved in dauer formation, are outside the focus of this thesis and so will not be discussed further.

Figure 1.6 Genetic pathway for dauer larva formation



This a representation of the wild-type genetic interactions between the three different pathways involved in dauer formation. Daf-c genes form dauers constitutively while Daf-d genes are dauer defective. The *daf-11* branch of the pathway represents cGMP signalling, while the *daf-1* branch represents TGF-β pathway. These to pathways integrate at *daf-12*, which encodes a nuclear hormone receptor (Antebi *et al.*, 2000). The *daf-2*, *age-1* and *daf-16* branch represents the IIS pathway and is discussed below. Briefly *daf-2* and *age-1* negatively regulate *daf-16*, promoting reproductive development, when *daf-16* is active it promotes dauer development or in adults extended longevity. Image taken from (Riddle and Albert, 1997).

1.3.5 Role of IIS in dauer formation and lifespan

The IIS pathway not only regulates dauer formation but also has a role in the regulation of lifespan in *C. elegans*. Hypomorphic mutations in several genes within the pathway have been shown to increase the adult lifespan. The IIS pathway (Figure 1.7) currently consists of 39 genes that encode insulin-like peptides, *ins-1* to *ins-37* (Duret *et al.*, 1998; Gregoire *et al.*, 1998; Kawano *et al.*, 2000; Kawano *et al.*, 2003; Pierce *et al.*, 2001),

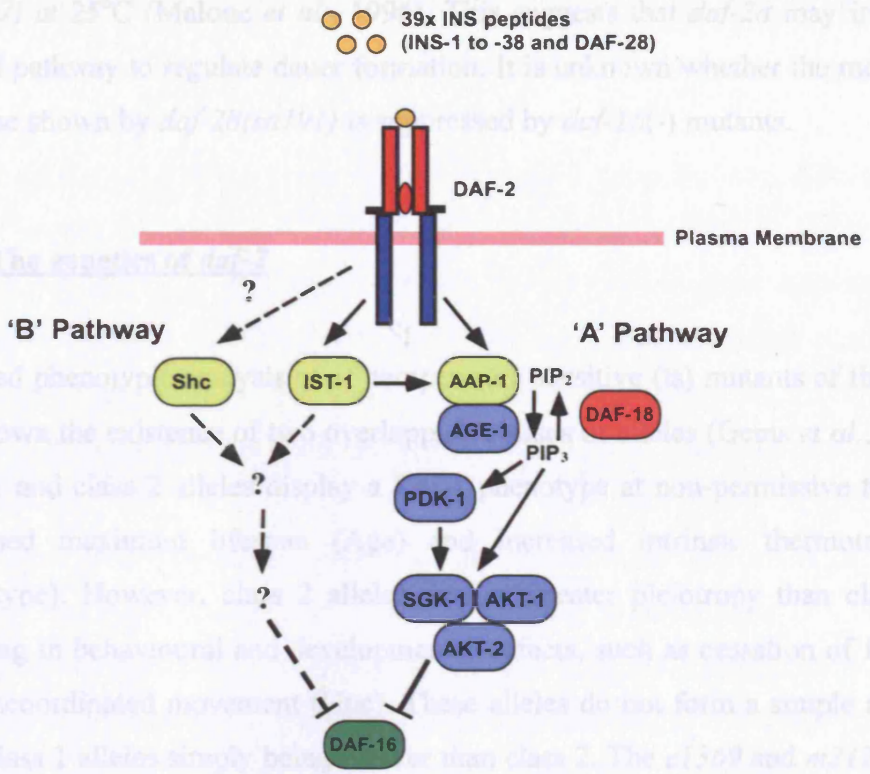
ins-38 (M. Costa, personal communication) and *daf-28* (Li *et al.*, 2003) (see Figure 1.8). The products of these genes are all predicted to be ligands, either agonists or antagonists (see chapter 7), for the DAF-2 receptor. The only one of the insulin-like peptide genes that has been shown to have a mutant phenotype is *daf-28*. The *daf-28(sa191)* allele is a semi-dominant allele that has a Daf-c phenotype at all temperatures, although its severity increases with temperature (Li *et al.*, 2003; Malone *et al.*, 1996; Malone and Thomas, 1994). This mutation causes a 12-13% increase in lifespan in mutants compared to wild type worms.

The *daf-2* gene encodes a receptor tyrosine kinase that is a member of the IR superfamily of proteins (for more details about *daf-2* gene structure and protein structure see chapter 4) (Kimura *et al.*, 1997). Hypomorphic mutations of *daf-2* have been shown to extend lifespan at least two-fold (Kenyon *et al.*, 1993). Activation of the DAF-2 protein leads to recruitment and activation of the phosphatidylinositol-3-OH kinase (PI 3-kinase) complex.

The PI 3-kinase complex in *C. elegans* consists of two different proteins. The regulatory subunit is encoded by *aap-1* (Wolkow *et al.*, 2002) and the catalytic subunit of the complex is encoded by *age-1* (Morris *et al.*, 1996). *age-1* was the first gene in *C. elegans* that was reported to extend lifespan (Friedman and Johnson, 1988a). The PI 3-kinase complex may be recruited to DAF-2 either by IST-1, a homologue of the vertebrate insulin receptor substrate (IRS) and the *D. melanogaster* CHICO proteins. Or DAF-2 may directly recruit the PI 3-kinase complex by direct interaction with the AAP-1 subunit through its C-terminal extension (Wolkow *et al.*, 2002). The PI 3-kinase complex then phosphorylates PIP2 to PIP3, which acts as a secondary messenger. Levels of PIP3 can be attenuated by the DAF-18 lipid phosphatase, which is the *C. elegans* homologue of the PTEN human tumour suppressor (Gil *et al.*, 1999; Mihaylova *et al.*, 1999; Ogg *et al.*, 1997; Ogg and Ruvkun, 1998).

Elevated levels of PIP3 are thought to recruit the 3-phosphoinositide-dependent kinase-1 (PDK-1), a serine/ threonine kinase encoded by the *pdk-1* gene (Paradis *et al.*, 1999), to the plasma membrane where it co-localizes with a trimeric complex formed by AKT-1/-2 and SGK-1 (serum- and glucocorticoid-inducible kinase-1) (Hertweck *et al.*, 2004; Paradis and Ruvkun, 1998). PDK-1 phosphorylates and activates this complex, which in turn phosphorylates DAF-16. Hypomorphic mutations affecting the *pdk-1* and *sgk-1* genes increase lifespan (Hertweck *et al.*, 2004; Paradis *et al.*, 1999). Interestingly, null mutations of *akt-1*/ -2 do not extend lifespan, but do have a Daf-c phenotype.

Figure 1.7 IIS pathway in *C. elegans*



A representation of the interaction between components of the IIS pathway in *C. elegans*. There are 39 genes that encode insulin-like peptides in the *C. elegans* genome, some of these are thought to bind to and activate the DAF-2 insulin/ IGF-1 receptor. This leads to the phosphorylation of AAP-1 (p85 homologue) either directly through its interaction with the C-terminal extension of DAF-2 or through IST-1 (homologue of IRS-1). This activates AGE-1 (p110 homologue) and leads to production of PIP3. Elevated levels of PIP3 activate PDK-1 (3-phosphoinositide-dependent kinase-1), which in turn activates a trimeric complex of serine/ threonine kinases formed by AKT-1, AKT-2 (PKB homologues) and SGK-1 (serum- and glucocorticoid-inducible kinase-1). Elevated levels of PIP3 may also directly activate these kinases. This complex negatively regulates DAF-16 by phosphorylating it. PIP3 levels can be attenuated by DAF-18 (PTEN homologue). These different components make up the only genetically defined output of DAF-2 and might represent the 'A' pathway in the model of *daf-2* signalling put forward by Gems *et al.* (1998). No components of the proposed 'B' pathway have yet been identified.

The *daf-16* gene encodes a forkhead transcription factor (Lin *et al.*, 1997; Ogg *et al.*, 1997). The principal role of the DAF-2/ AGE-1 pathway is to antagonize DAF-16 activity, by phosphorylating it. When DAF-16 is phosphorylated and inactive it is localized to the cytoplasm of the cell. When DAF-16 is unphosphorylated and active it is localized to the cell nucleus (Henderson and Johnson, 2001). The lifespan extension and Daf-c phenotype caused by mutations in *daf-2*, *age-1*, *pdk-1* and *sgk-1* are all *daf-*

16 (+) dependent, and are fully suppressed by *daf-16* null mutants. Interestingly, the Daf-c phenotype of *daf-28(sa191)* is only partially suppressed by *daf-16(m26)* or *daf-16(m27)* at 25°C (Malone *et al.*, 1996). This suggests that *daf-28* may interact with a second pathway to regulate dauer formation. It is unknown whether the modest lifespan increase shown by *daf-28(sa191)* is suppressed by *daf-16*(-) mutants.

1.3.6 The genetics of *daf-2*

Detailed phenotypic analysis of 15 temperature sensitive (ts) mutants of the *daf-2* gene has shown the existence of two overlapping classes of alleles (Gems *et al.*, 1998). Both class 1 and class 2 alleles display a Daf-c phenotype at non-permissive temperatures, increased maximum lifespan (Age) and increased intrinsic thermotolerance (Itt phenotype). However, class 2 alleles display greater pleiotropy than class 1 alleles resulting in behavioural and developmental defects, such as cessation of feeding (Eat) and uncoordinated movement (Unc). These alleles do not form a simple allelic series, with class 1 alleles simply being weaker than class 2. The *e1369* and *m212* alleles have the most severe Daf-c phenotype of all ts alleles, forming 55% and 28% dauers, respectively, at 15°C, yet both are class 1 (Gems *et al.*, 1998).

The genetics of *daf-2* can become very confusing when using epistasis analysis to map genetic interactions. At least three different genes have been shown to have differing interactions with the different classes of *daf-2* allele. These are *daf-12* (Gems *et al.*, 1998; Larsen *et al.*, 1995), *rop-1* (Labbe *et al.*, 2000) and *daf-9* (Gerisch *et al.*, 2001). The *daf-12* gene encodes a nuclear hormone receptor that lies at the convergence of the dauer formation and heterochronic pathways (Antebi *et al.*, 2000). The hypomorphic *daf-12(m20)* allele is Daf-d, when a class 1 *daf-2* allele is placed into this background at 25°C, the *m20* allele suppresses the Daf-c phenotype of the *daf-2* allele and weakly suppresses the Age phenotype (Gems *et al.*, 1998; Larsen *et al.*, 1995). However, when a class 2 *daf-2* allele is placed into this background at 25°C, the *m20* allele causes an embryonic/ L1 arrest phenotype and enhances the Age phenotype of the *daf-2* allele .

The *rop-1* gene encodes an RNA quality-control pathway component that functions in the dauer formation pathways (Labbe *et al.*, 2000). The *rop-1(pk93)* allele enhances the Daf-c phenotype of *daf-2(e1370)* and *daf-2(m579)* (both class 2) at 22.5°C

and suppresses the Daf-c phenotype of *m41* (class 1) and *m596* (originally labelled a class 2 allele in Gems *et al.*, 1998, but reclassified as class 1 in this project). The *daf-9* gene encodes a cytochrome P450 (Gerisch *et al.*, 2001; Jia *et al.*, 2002). The *daf-9(rh50)* enhances the Age phenotype of *daf-2(e1370)* and weakly suppresses Age in *daf-2(e1368)* (class 1) (Gerisch *et al.*, 2001).

Gems *et al.* (1998) proposed that the *daf-2* gene has two different activities termed *daf-2A* and *daf-2B*. Class 1 alleles are *daf-2A (-) daf-2B (+),* while class 2 mutants are *daf-2A(-) daf-2B(-)*. The *daf-2A* and *daf-2B* gene activities may each represent an output, from the DAF-2 receptor, for parallel branches controlling dauer formation (A and B pathways). These pathways must converge on the *daf-16* gene, as *daf-16(-)* mutants suppress the pleiotropic effects of both classes of *daf-2* allele (Gems *et al.*, 1998).

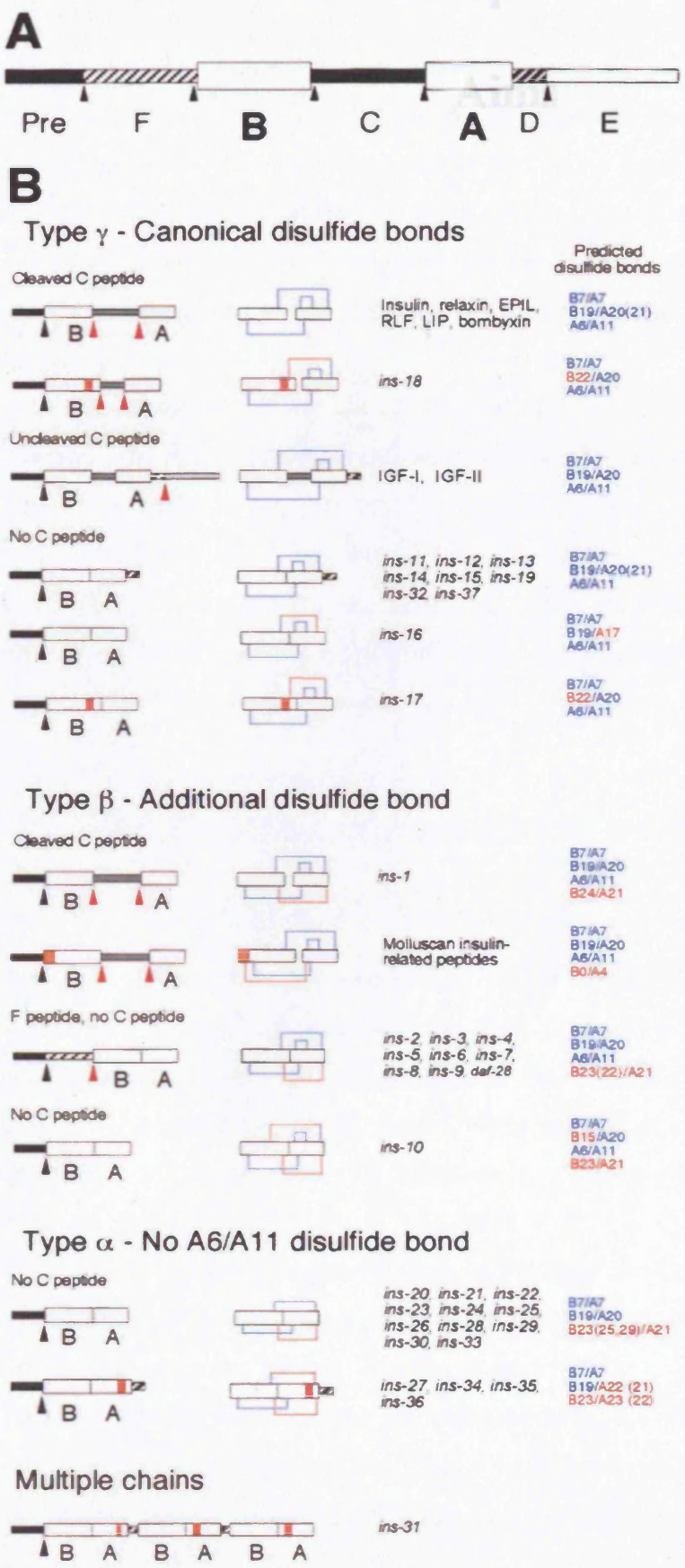
There is increasing evidence for the existence of at least a second parallel pathway that signals through the DAF-2 receptor to antagonise DAF-16, which is independent of the AGE-1 PI 3-kinase pathway. Firstly, activating mutants of both *pdk-1* and *akt-1* are able to fully suppress the Daf-c phenotype of the *age-1(mg44)* allele, but only partially suppress Daf-c in the *daf-2(e1370)* allele (Paradis *et al.*, 1999; Paradis and Ruvkun, 1998), even though *age-1(mg44)* is a non-conditional allele whereas *daf-2(e1370)* is ts. This suggests that the DAF-2 receptor is able to antagonise DAF-16 through a second pathway that is independent of AGE-1 (see Figure 1.7). This second pathway may involve *ist-1*, as RNAi of *ist-1* in *age-1(mg44); akt-1(mg144gf)* double mutants causes a large proportion of the animals to arrest as dauers (Wolkow *et al.*, 2002). Shc/ RAS signalling may also be involved in the second pathway (see Chapter 5). Recently it has been shown that LET-60 (RAS homologue in *C. elegans*) modulates IIS and has a weak role in dauer formation and ageing (Nanji *et al.*, 2005). Thus, *daf-2*A activity may involve the AGE-1/PI 3-kinase pathway and *daf-2*B activity may involve the RAS pathway.

The existence of so many putative ligands (39 currently) for the DAF-2 receptor also implies that there might be more than one output from the receptor (Duret *et al.*, 1998; Gregoire *et al.*, 1998; Kawano *et al.*, 2000; Kawano *et al.*, 2003; Li *et al.*, 2003; Pierce *et al.*, 2001) (M. Costa, personal communication). Two of these insulin-like genes, *ins-1* and *ins-18* as well as human insulin, actually antagonise the DAF-2 receptor at high gene dosage and enhance dauer formation in both wild type worms and weak *daf-2* mutants. Overexpression of *ins-9, ins-22* and an *ins-19/-31* hybrid gene

have no effect on dauer formation in either wild-type worms or weak *daf-2* mutants (Pierce *et al.*, 2001).

The existence of multiple insulin-like ligands that either activate or antagonise the DAF-2 receptor could provide support for a model put forward to explain the differential effects of somatic gonad and germline ablations on life span in class 1 and class 2 *daf-2* alleles (Hsin and Kenyon, 1999). The somatic gonad in *C. elegans* generates a signal that acts to increase maximum life span and *daf-2* activity is required for this effect. The germline, however, produces a signal, dependent on *daf-16* and *daf-12*, which shortens life span in an exactly opposing measure to the somatic gonad. Ablation of the germline precursors in both classes of *daf-2* allele leads to an enhancement of the Age phenotype. Whereas ablation of the somatic gonad precursors in class 2 alleles leads to an enhancement of the Age phenotype, ablation in class 1 alleles has little effect on *daf-2(e1368)* and only a mild effect on the *daf-2(m41)* Age phenotype.

Figure 1.8 Molecular architecture of insulin-like peptides



Panel **A** shows the amalgamated domain structure of members of the insulin superfamily. All members have at least a signal peptide (Pre) and a B and A chain, all non-*C. elegans* peptides also have a cleaved C peptide. IGF-1 and IGF-II have D and E domains at their C-terminus. Certain members of the *C. elegans* β-type INS peptides have an N-terminal F-peptide following their signal peptide.

Panel **B** shows the specific architecture of each of the three different sub-types of *C. elegans* INS peptides (γ, β and α) and the location of the cysteines predicted to be involved in disulphide bonds in the B and A chains. The peptides marked with red boxes and lines have predicted disulphide bonds that are additional to the canonical bonds found in vertebrate members of the insulin superfamily. Canonical disulphide bonds are indicated by blue lines. Image adapted from (Hertweck *et al.*, 2004;

Chapter 2

Aims

2.1 Aims of this thesis project

The overall aim of this thesis is to further our understanding of the *daf-2* gene in *C. elegans* and its role in dauer formation and longevity. Specifically this thesis seeks to understand the nature of the two different phenotypic classes formed by ts alleles of *daf-2* (Gems *et al.*, 1998), as well as the functional role of the large N- and C-terminal extensions found in the DAF-2 protein (Kimura *et al.*, 1997). This thesis also probes the role of two putative ligands of the DAF-2 receptor, INS-7 and INS-35, in determining longevity.

The structure of the results chapters of this thesis is as follows; chapter 4 compares the *daf-2* gene in *C. elegans* to its homologues in other nematodes. Chapter 5 presents a sequence analysis of different *daf-2* alleles. Chapter 6 tests some of the predictions about *daf-2* function generated by chapter 5. Chapter 7 moves away from the DAF-2 receptor and instead focuses on two of its putative ligands and their role in longevity. Below I set out a brief rationale behind each of the results chapters and the approaches used to answer these questions.

2.2 Chapter 4 Identifying *daf-2* homologues in the nematodes *C. briggsae* and *B. malayi*

The *daf-2* gene encodes the sole member of the insulin/ IGF-1 receptor protein family within the *C. elegans* genome (Kimura *et al.*, 1997). This was the only member of this gene family to have been characterised in the nematode phylum at the start of this project. The predicted DAF-2 protein, while possessing all the defining domains and features of a member of the IR superfamily, also possesses large N-terminal and C-terminal extension (see Figure 4.3) (Kimura *et al.*, 1997). The N-terminal extension is novel and has no homology to any known proteins. The C-terminal extension is similar to that found in the *D. melanogaster* dInR protein (Fernandez *et al.*, 1995; Ruan *et al.*, 1995), and contains several tyrosine residues that be phosphorylated upon ligand binding. Three of these putative phosphotyrosines may serve as direct docking sites for the *C. elegans* homologues of p85, PLC-γ, SHP-2 and Grb2 (Kimura *et al.*, 1997). Thus, the C-terminal extension in DAF-2 may serve a similar function to the IRS proteins in

vertebrate IIS pathways, even though the *C. elegans* genome does contain a gene encoding an IRS homologue known as IST-1, a known component of the IIS pathway (Wolkow *et al.*, 2002).

Several nematode genomes have been or are being sequenced since this project began. This enables rapid identification of *daf-2* homologues within these species and allows us to ask the following questions; are the N- and C-terminal extensions seen in DAF-2 unique to *C. elegans*? Or do they have a broader phylogenetic distribution? Which might imply an important role for these extensions in IIS pathway activity in nematodes. Currently the *daf-2* promoter remains undefined, can additional *daf-2* gene orthologue sequences help define regulatory elements in *C. elegans*? There is tremendous variation in lifespans across the nematode phylum, ranging from three days in free-living *Strongyloides ratti* adults to 15 years in the parasite *Loa loa* (Gardner *et al.*, 2004; Gems, 2000). Can some of the differences in lifespan between short-lived and long-lived nematodes be due to differences in their DAF-2 receptors, given that DAF-2 is a known gerontogene?

I have used several different approaches to address these questions. The *Caenorhabditis briggsae* and *Brugia malayi* genomes have been sequenced (Ghedin *et al.*, 2004; Stein *et al.*, 2003) and the *Caenorhabditis remanei* genome is currently being sequenced. *C. briggsae* and *C. remanei* are closely related to *C. elegans* and have similar lifespans (McCulloch and Gems, 2003). *B. malayi* is a long-lived human parasitic nematode that has a lifespan of at least 6-7 years (Wang *et al.*, 1994). I have sequenced the cDNA of the *daf-2* homologues from *C. briggsae* and *B. malayi* and modelled the gene structure of *daf-2* in *C. remanei*. Using these sequences I have performed a comparative analysis of both the *daf-2* gene and the predicted DAF-2 proteins from all four nematode species.

2.3 Chapter 5 Sequence analysis of *daf-2* mutant alleles

Temperature-sensitive mutations within the *daf-2* gene do not fall into a simple allelic series, but instead can be separated into two different phenotypic classes (1 and 2) (Gems *et al.*, 1998). Both classes display Daf-c, Age and Itt phenotypes at non-permissive temperatures. Class 2 alleles, however, also display additional pleiotropic phenotypes. These classes also differ in their epistatic interactions with other genes

involved in dauer formation and often have opposite phenotypes (Gems *et al.*, 1998; Gerisch *et al.*, 2001; Labbe *et al.*, 2000; Larsen *et al.*, 1995). Currently, there is no understanding of the molecular nature of the difference between the two classes. However, sequencing of several class 1 and 2 alleles by Kimura *et al.* (1997) suggests that the different classes may correlate with mutations in particular regions of the DAF-2 protein (Gems *et al.*, 1998). A formal model for the differences in genetic interactions between the two classes posits that the *daf-2* gene has two activities *daf-2A* and *daf-2B*, which may correspond to different outputs from the DAF-2 receptor ('A' and 'B' pathway, see Figure 1.7) (Gems *et al.*, 1998).

The human insulin and IGF-1 receptors have been well characterised both biochemically and structurally (reviewed in De Meyts and Whittaker, 2002). Given that DAF-2 is homologue of *h*IR and *h*IGF-1R, can the data available for these two receptors be applied to DAF-2 in order to generate hypotheses about the molecular nature of the two classes of *daf-2* allele? Mutations in *h*IR can be classified into 5 types depending on the point at which they affect receptor function (see Figure 1.4) (Taylor *et al.*, 1992). Can mutations in DAF-2 be classified using the Taylor *et al.* (1992) scheme? *h*IR and *h*IGF-1R both have multiple signalling outputs, such as PI 3-kinase, CAP/ Cbl and Ras/ MAP kinase pathways (Saltiel and Kahn, 2001). Are the different classes of *daf-2* allele linked to mutations in DAF-2 that alter its signalling outputs? The *C. elegans* genome contains 39 genes that encode insulin-like peptides (Duret *et al.*, 1998; Gregoire *et al.*, 1998; Kawano *et al.*, 2000; Kawano *et al.*, 2003; Li *et al.*, 2003; Pierce *et al.*, 2001; M. Costa, personal communication), yet *daf-2* is the only gene encoding insulin/ IGF-1-like receptor. Does the structure of DAF-2 indicate why there are so many ligands for it?

To address these questions I have identified the lesions in several class 1 and 2 alleles and compared the location of these mutations across the receptor to previously published alleles. Using this expanded repertoire of *daf-2* mutations I have carried out an integrated analysis of DAF-2 that combines homology modelling and data mining of *h*IR and *h*IGF-1R mutants to attempt to predict the molecular effects of each mutation on DAF-2 function.

2.4 Chapter 6 Further analysis of *daf-2* alleles

The majority of this chapter deals experiments designed to test predictions about DAF-2 function generated by the results of chapter 5 and so will not be discussed here. However, part of this chapter also deals with the question of the *daf-2* null phenotype. Currently, the *m65* allele is the only candidate for a *daf-2* null based on epistasis data with *daf-16* (Larsen *et al.*, 1995; M. Nanji and D. Gems, unpublished results). Recently, the Japanese *C. elegans* knockout consortium generated a deletion allele of *daf-2* at our request. In this chapter I compare the Daf-c phenotypes of the *m65* and deletion alleles, both as single mutants and as *daf-16*; *daf-2* double mutants, to determine if either or both represent a *daf-2* null.

2.5 Chapter 7 Analysis of the *ins-7* and *ins-35* genes encoding putative ligands for DAF-2

This final results chapter deals with two of the 39 genes encoding insulin-like peptides in the *C. elegans* genome. McElwee *et al.* (2004) performed microarray experiments comparing the genome-wide expression profiles of *glp-4*; *daf-2* double mutants to *glp-4 daf-16*; *daf-2* triple mutants using the Affymetrix™ *C. elegans* oligonucleotide array. Of the 39 insulin-like peptide genes in the genome only 23 are represented on this array platform. Nine of these *ins* genes show statistically significant alterations in expression levels between the *glp-4*; *daf-2* double mutants and the *glp-4 daf-16*; *daf-2* triple mutants (Table 7.1.1). The *ins-7* and *ins-35* genes show the greatest changes in expression between these two genotypes, with *ins-7* being down-regulated by ~75% and *ins-35* up-regulated ~24-fold, respectively, in *glp-4*; *daf-2* mutants compared to *glp-4 daf-16*; *daf-2* triple mutants.

The current favoured hypothesis for the existence of so many insulin-like peptides in *C. elegans* suggests that some of these proteins behave like classical agonists and activate the DAF-2 receptor, whilst others may function as antagonists and inhibit the receptor (Pierce *et al.*, 2001). Potentially *ins-7* may be acting an agonist of DAF-2, which is supported by the fact the RNAi of this gene has been reported to extend lifespan (Murphy *et al.*, 2003). Conversely, *ins-35* may be an antagonist of DAF-2,

based on its up-regulation in the *glp-4*; *daf-2* background. If *ins-35* is an antagonist then it could represent the first example of a *C. elegans* longevity hormone. RNAi of *ins-35* in *daf-2* mutant background can test for this possibility, as knockdown of this gene would be predicted to suppress the extended lifespan of *daf-2(-)* animals.

Chapter 3

General materials and methods

3.1 *C. elegans* strains and culture

3.1.1 Strains

The N2 *Caenorhabditis* Genetics Center (CGC) male stock was used as the wild-type strain throughout this thesis, as the lifespan of this strain most closely resembles that of the original N2 isolate (Gems and Riddle, 2000). All other strains used throughout this thesis are listed in each chapter. Unless stated otherwise, all strains used were from the Gems lab strain collection or constructed by myself.

3.1.2 Routine culture of stocks

All stocks were grown in 60mm tissue culture plates (Greiner Bioscience) containing nematode growth medium (NGM) agar seeded with *E. coli* strain OP50 (Brenner, 1974). Stocks were generally maintained at 15°C unless otherwise stated. NGM agar is uracil-limited and the OP50 strain is a uracil auxotroph, therefore the NGM agar prevents overgrowth of the bacterial lawn. This stops the lawn from obscuring the worms when viewed under a microscope. Animals were inspected using a Nikon SMZ645 dissecting scope.

Animals were transferred to fresh plates at least once a week using a thin platinum wire (worm pick). The worm pick was always sterilised by flaming before being allowed to cool down. Once cooled the pick was coated with fresh OP50 and used to pick up animals for transfer. The worm was then gently set down on the bacterial lawn of the fresh plate and worms were allowed to crawl away. For stocks maintained as hermaphrodite only cultures 3-5 gravid adults were transferred to fresh plates. For stocks maintained as mixed male and hermaphrodite cultures 6-10 L4 males were transferred to fresh plates along with 3 L4 hermaphrodites. Males were initially generated using heat shock. Six to ten L4 hermaphrodites were exposed to 30°C for 6 hours before being transferred back to 15°C. This procedure increases the incidence of X chromosome non-disjunction during meiosis and thus increases the rate of

spontaneous male formation. Males were sought amongst the progeny of these worms, and then mated with hermaphrodites, as described above, to maintain their production.

3.1.3 Axenisation of stocks

Axenisation is an alkaline hypochlorite based method that allows fertilised eggs to be retrieved from gravid adults. This method can be used to clean up stocks that have become contaminated with non-OP50 bacteria or yeast. This method is also essential in removing OP50 cells from eggs that are being transferred onto HT115(DE3) cells for RNAi. This procedure is also the first step in growing stage synchronous cultures of strains.

Procedure and reagents

1. Gravid adults were washed of plates using a total volume of 6 ml of M9 buffer (Brenner, 1974) and into a sterile 15 ml centrifuge tube. The agar plates will retain some of the buffer, so the final volume in the tube should be ~5 ml.

2. The animals were then incubated with 2 ml of 5% (Volume/Volume) NaOCl solution and 1 ml of 5M KOH for three minutes. This treatment causes the gravid hermaphrodites to lyse and release their fertilised eggs, as well as killing any contaminants.

3. After the alkaline hypochlorite incubation worms were centrifuged at 1000 rpm for 1 min. The supernatant was drawn off leaving behind a pellet of free eggs and lysed worms.

4. The pellet was resuspended in 10 ml of M9 and then centrifuged again at 1000 rpm. The supernatant was removed and the pellet resuspended in another 10 ml of M9 and centrifuged a further two times.

5. After the final centrifugation, only 9.5 ml of the supernatant was removed. The pellet was resuspended in the remaining 0.5 ml of M9. 200 µl of the egg

suspension were then aliquoted on to fresh plates seeded with OP50. The suspension was deliberately placed away from the bacterial lawn. This ensured that larvae crawled away from the lysed bodies of the parental worms.

6. These plates were stored overnight at 15°C. The next day the hatched larvae that were feeding on the bacterial lawn were transferred to new plates and the plates with the lysed worms were discarded.

3.1.4 Stage synchronous cultures

Stage synchronous cultures of worms are essential for certain types of experiment. For example, in this thesis I needed stage synchronous cultures for the quantitative PCR analysis of *daf-2* transcript levels in the *daf-2(e1369)* allele to ensure that any difference measured was not due to developmental differences. Stage synchronous cultures of worms can be generated using the axenisation protocol listed above. However, instead of performing the last two steps listed in the protocol the pellet from the final wash is resuspended in 10 ml of M9. This suspension is transferred into a sterile 25 ml conical flask and left in a shaking incubator, rotating at 200 rpm, at 20°C for 24 hours. This procedure allows the eggs to hatch into L1 larvae. However, due to the absence of a food source all the larvae arrest in the L1 diapausal state.

The arrested L1 larvae can be harvested by transferring the M9 suspension from the conical flask into a sterile 15 ml centrifuge tube. The tube is then centrifuged at 1000 rpm. The supernatant can then be removed and the worms resuspended in the appropriate volume of M9 that is suitable for the experiment. The resuspended worms can be placed onto seeded plates, either OP50 for standard culture or HT115(DE3) for RNAi.

3.2 Molecular biology in *C. elegans*

3.2.1 RNA extraction

All total RNA extractions from both *C. elegans* and *C. briggsae* were carried out using TRIzol reagent (Invitrogen) and following the manufacturer's instructions. Two slightly different methods were used depending on the purpose of the RNA sample. RNA samples for the synthesis and sequencing of cDNA from different *daf-2* alleles and the *Cb-daf-2* gene (Chapter 4) were smaller scale preparations extracted without the use of a ribolyser. RNA used in the *daf-2(e1369)* QPCR experiments was extracted with the use of a ribolyser, which increases yield, and were larger scale preparations. Both methods are outlined below.

Procedure and reagents

Small-scale extractions

1. For the small-scale RNA extractions, worms were grown in mixed stage cultures at 15°C on six to ten 60mm tissue culture plates per strain. When the plates reached high-population density and were close to exhausting the OP50 lawn the animals were washed from the plates into a 15 ml centrifuge tube using 10 ml of M9. The plates were grown to high-density to maximise the amount of nematode RNA recovered and minimise the amount of bacterial RNA recovered.

2. The worms were then centrifuged at 1000 rpm to pellet the worms and the supernatant removed. The worms were then resuspended in 200 µl of M9, which was then transferred into a 1.5 ml centrifuge tube (RNase and DNase-free, Anachem).

3. 800 µl of TRIzol reagent (Invitrogen) was then added to the tube and vortexed for 30 min at room temperature (RT) and then left to incubate for a further 15 min at RT.

4. The worm-TRIzol mixture was then centrifuged at 14,000 rpm at 4°C for 10 min to deposit the insoluble worm cuticles. The supernatant was then transferred to a new 1.5 ml centrifuge tube. The rest of the TRIzol extraction was then carried out according to the manufacturer's instructions.

5. The RNA pellet formed by the extraction process was then air-dried at 65°C, before being dissolved in 50 µl of RNase-free water (Invitrogen). 5 µl of the dissolved RNA was diluted into 1 ml of RNase-free water (1:200 dilution) and was placed in a spectrophotometer to measure the quantity of RNA produced by the extraction. The rest of the sample was stored at -80°C until required.

Large-scale extractions

1. For the large-scale RNA extractions used in the QPCR experiments worms were grown in synchronous culture (section 3.1.4). Worms were grown on four 100mm plates per strain at 15°C until they reached the L4 stage. They were then shifted to 25°C overnight, before they were washed into 15 ml centrifuge tubes using 10 ml of M9.

2. The worms were then centrifuged at 1000 rpm for 1 min to pellet the animals. The supernatant was removed and the worms were resuspended in 1 ml of TRIzol reagent. The TRIzol-worm suspension was transferred into a Lysing Matrix D tube (Q-biogene) and frozen at -80°C overnight. This was done to help disrupt the worm cuticle and increase yield of RNA

3. The Lysing Matrix D tube was then thawed out at RT. Once thawed the tube was placed in a FP220 ribolyser (Q-biogene) and subjected to two 30 second cycles of homogenisation at a speed setting of 4.5 and then placed on ice.

4. 200 µl of chloroform ($CHCl_3$) was then added to the homogenate and the mixture was vortexed for 15 seconds before being left to incubate at RT for 3 min.

5. The homogenate-chloroform mixture was then centrifuged at 14,000 rpm at 4°C for 15 min to deposit the insoluble worm cuticles. The aqueous layer of supernatant was then transferred to a new 1.5 ml centrifuge tube. The rest of the TRIzol extraction was then carried out according to the manufacturer's instructions.

6. Drying of the RNA pellet and quantification of the RNA yield was carried as described above for the smaller scale preparations.

3.2.2 First strand cDNA synthesis

Single-stranded cDNA was synthesised from RNA samples generally using an oligo-δT primer. However, for the majority of the *daf-2* allele sequencing work I used a *daf-2* gene-specific primer (see chapter 5), except in the case of *daf-2(e1369)*, where I used both. The first strand cDNA synthesis was carried out using the Reverse-iT first strand cDNA synthesis kit (ABgene) and the protocol is outlined below.

Procedure and reagents

1. 5 µg of total RNA were added to a low profile RNase/ DNase-free PCR tube (ABgene). 1 µl of oligo-δT primer or 1 µl of gene-specific primer was then added to the tube. RNase-free water (Invitrogen) was then added to the tube to take the volume of this mixture to 13 µl. The volume of water added was dependent on the concentration of the RNA. The mixture was then heated to 65°C and then placed on ice. This step helps eliminate secondary structure in mRNA that might inhibit reverse transcription.

2. The following reagents were then added to the PCR tube: 4 µl of first strand cDNA buffer, 2 µl of dNTPs and finally 1 µl of the Reverse-iTase (ABgene proprietary reverse transcriptase). The mixture was gently mixed and then spun down in a centrifuge at 800 rpm for 1 min.

3. The PCR tube was then placed in a GeneAmp 2700 thermal cycler (Applied Biosystems). The following programme was used: 47°C for 60 min for the reverse transcription, 10 min at 75°C to denature the Reverse-iTase followed by a 4°C hold. Once completed the first strand cDNA was stored at -20°C until needed.

3.2.3 Genomic DNA extraction

Genomic DNA (gDNA) was needed for several different purposes, such as confirming mutations found in the cDNA of *daf-2* alleles (chapter 5), following deletion alleles during strain construction (chapter 6) and sequencing of the upstream DNA and first intron from *daf-2(e1369)* (chapter 5). In most cases genomic DNA was isolated from 5-10 adult worms, but can be extracted from just single worms when confirming genotypes. The protocol for multiple worms is outlined below, the procedure for single worm lysis can be found in the single worm PCR section 3.2.4.

Procedures and reagents

1. The worms were first placed in 10 µl of filter-sterilised lysis buffer in a PCR tube. This is made up in RNase-free water and has the final composition;

   50mM KCl
   10mM Tris (pH8.3)
   2.5mM $MgCl_2$
   0.45% Nonidet P-40
   0.45% Tween 20
   0.01% gelatin
   60 µg/ml proteinase K

2. The worms were then frozen at -80°C for 10 min to help break up the cuticle. The PCR tube was placed in a thermal cycler. The following programme was used: 65°C for 60 min, 15 min at 95°C to denature the proteinase K followed by a 4°C hold. Once completed the genomic DNA was stored at -20°C until needed.

3.2.4 PCR protocols

I have used several different types of PCR protocol in this thesis. For the amplification of cDNA and gDNA fragments for use in sequencing I usually used a standard PCR protocol. For difficult templates, colony PCR from bacterial cells (chapter 7) or for single worm PCR for genotyping (chapter 6) I employed touchdown PCR. All PCRs were carried out using 2x ReddyMix PCR master mix (ABgene) containing 1.5mM $MgCl_2$. The different methods are outlined below.

Procedures and reagents

Standard PCR

1. Generally 1 µl of template DNA (either cDNA or gDNA) was added to 12 µl of RNase/ DNase free water in a PCR tube. 1 µl of each relevant (5' and 3') primer was then added to the tube to give a volume of 15 µl. 15 µl of the 2x ReddyMix PCR master mix was then added into the tube and gently mixed by pipetting.

2. The PCR tube was then place in the thermal cycler and the following program was used:

   94°C for 5 min

   Followed by 35 cycles of
   94°C for 30 seconds,
   55°C for 30 seconds,
   72°C for X minutes (see Table 3.2.1)

   4°C final hold

   Once the PCR was completed, 10 µl of the sample was visualised on an agarose gel to confirm amplification and the rest was stored at -20°C until needed.

Touchdown PCR

In a touchdown PCR the annealing temperature decreases by 1°C every cycle for a user-defined number of cycles. This procedure can help eliminate non-specific primer binding so that there is better amplification of the desired fragment.

1. Generally 1 µl of template DNA (either cDNA or gDNA) was added to 12 µl of RNase/ DNase free water in a PCR tube. In the case of colony PCR I resuspended some bacterial cells sampled with my worm pick in 12 µl of water. 1 µl of each relevant (5' and 3') primer was then added to the tube to give a volume of 15 µl. 15 µl of the 2x ReddyMix PCR master mix was then added into the tube and gently mixed by pipetting.

2. The PCR tube was then place in the thermal cycler and the following program was used:

   94°C for 5 min

   Followed by 10 cycles of
   94°C for 30 seconds,
   65°C* for 30 seconds (*decreases by 1°C every cycle for first 10 cycles)
   72°C for X minutes (see Table 3.2.1)

   Followed by 25 cycles of
   94°C for 30 seconds,
   55°C for 30 seconds,
   72°C for X minutes (see Table 3.2.1)

   4°C final hold

   Once the PCR was completed, 10 µl of the sample was visualised on an agarose gel to confirm amplification and the rest was stored at -20°C until needed.

Single worm PCR

For single worm PCR I always used touchdown PCR. Single worms were placed in 2.5 µl of lysis buffer (section 3.2.3) in a PCR tube. The tube was then frozen for 10 min at -80°C to help disrupt the cuticle of the worm. The tube was then placed in a thermal cycler and the following programme was used: 65°C for 60 min, 15 min at 95°C to denature the proteinase K followed by a 4°C hold. The tube was removed from the cycler and kept on ice. 1 µl of relevant primer was added to the lysate followed by 5.5 µl of RNase/ DNase free water to bring the volume in the PCR tube to 10 µl. Finally 10 µl of 2x ReddyMix PCR master mix was then added into the tube and gently mixed by pipetting. The PCR tube was then placed into a thermal cycler and the touchdown PCR protocol described above was used. Once the PCR cycle was complete the entire reaction mixture was run out on an agarose gel for visualisation.

Quantitative PCR (QPCR)

All QPCR reactions were carried out in an ABI PRISM 7000 sequence detection system. I used 2x Absolute SYBR green ROX PCR master mix (ABgene) for the reactions, following the manufacturer's instruction. The final volume of the reaction in each well was 25 µl. All reactions were set up according to plate layout given in Chapters 6 and 7. The *ama-1* primers were both used at a final concentration of 50nM. The *daf-2* primers were used at a final concentration of 300nM for F2F and 50nM for F1R. The *ins-7* primers were used at a final concentration of 300nM for *ins-7*Fwd and 50nM for *ins-7*Rev. The *ins-35* primers were used at final concentration of 300nM for both primers. The relevant cDNA template was added in 2 µl aliquots to each reaction, except for 2x and 4x N2 cDNA samples, which were added in 4 µl and 8 µl aliquots (using a 1x stock) to each reaction (see plate layouts). Reactions were set up in a 96-well Thermo-Fast PCR plate (ABgene) and sealed with ABsolute QPCR optically clear adhesive film. The PCR protocol for the ABI PRISM 7000 sequence detection system was as follows: 50°C for 2 min, 95°C for 15 min (denatures antibody bound to Taq) then 40 cycles of 95°C for 15sec, 55°C for 30sec, 72°C for 30sec. The reactions were always run with the dissociation protocol switched on.

Table 3.2.1 Extension times used in all PCR protocols

Size of DNA fragment to be amplified	< 0.5Kb	< 1Kb	< 2Kb	< 3Kb
Extension time to be used (X)	30 seconds	1 minute	2 minute	3 minute

3.2.5 Electrophoresis of DNA fragments in agarose gels

All PCR products were run agarose gels to ensure successful PCR and amplification of the correct product. Similarly, the products of all restriction digests (see chapter 7) were checked to ensure they were correct by running the product through an agarose gel. For DNA products less than 1Kb in length I used a 1% gel (Weight/Volume) made up with 1x TAE buffer, for products greater than 1Kb I used a 0.8% gel. All gels contained 2 µl of ethidium bromide per 50 ml of agarose gel. Most samples were run in 10 µl volumes, except for DNA from single worm PCR and restriction digests, in which the entire reaction volume was run. All samples were run with 10 µl of ReddyRun 1Kb DNA ladder (ABgene) to aid identification of fragment sizes. Bands were then visualised using a UV transilluminator. Images of the bands were then captured using a digital camera and saved as tiff files, for future reference.

3.2.6 Sequencing of DNA fragments

Sequencing of PCR amplified DNA was used in obtaining the cDNA sequences of the *Cb-daf-2* and *Bm-daf-2* gene (chapter 4), locating and confirming the site of mutations in *daf-2* alleles (chapter 5), confirming the identity of *daf-2*; *daf-28* double mutants (chapter 6) and confirming the identity of RNAi constructs (chapter 7). Sequencing reactions were carried out using Big Dye Termination mix version 2.0 (Applied Biosystems). Given the volume of sequencing carried out in my project, I used a dilution protocol that allows good quality sequencing reads to be achieved from just an eighth of the normal quantity of Big Dye, thereby achieving 1600 reactions from a 200 reaction kit.

Procedures and reagents

1. First, the DNA to be sequenced was purified form the PCR mixture. This was done in two different ways. For fragments that had amplified only a single product the DNA was precipitated from the PCR master mix using the Microclean reagent (Microzone) and following the manufacturer's instructions. For PCR reactions that generated multiple bands the correct band (determined by size) was extracted from the gel and purified using the Qiaquick gel extraction kit (Qiagen) and following the manufacturer's instructions.

2. Once purified the DNA was resuspended in 30 µl of RNase/ DNase free water. 5.5 µl of the DNA sample was then added to a PCR tube. The following reagents were then added to the tube: 5 µl of Better Buffer (Microzone), 1 µl of Big Dye Termination mix version 2.0 (Applied Biosystems) and 2 µl of RNase/ DNase free water. Finally, 1.5 µl of the relevant primer was added to the mix (5' primer for a + strand read or the 3' primer for a – strand read).

3. The tube was then place in a thermal cycler and the following programme was used;

   25 cycles of

   96°C for 10 seconds
   55°C for 5 seconds
   60°C for 4 minutes
   4°C final hold

4. Once the sequencing reaction was complete the sample was removed from the cycler transferred from the PCR into a 1.5 ml centrifuge tube and the DNA precipitated by incubation with 80 µl of 80% isopropanol for 10 min at RT followed by centrifugation 13000 rpm for 10 min.

5. The supernatant was discarded and the pellet washed with 150 µl with 70% isopropanol before being centrifuged again at 13000 rpm for 10 min. Again the supernatant was discarded and sample was air-dried at 65°C.

6. The sample was then mixed with 10 µl of high purity formamide mix and heated at 65°C for 5 min. Samples were then transferred to 96-well plates and then heated for a further 4 min at 96°C and then cooled on ice for 5 min before being loaded into an ABI PRISM 3100 Genetic Analyzer™ (Applied Biosystems). Alice Smith and Mari Wyn Burley, the Goldstein lab sequencing technician and lab manager, respectively, carried out this final step and then kindly emailed the reads upon completion of the run to me.

Not all sequencing reactions were carried out as above. Towards the end of my project, samples were sequenced using the DNA sequencing service provided by the Wolfson Institute for Biomedical Research at UCL (http://www.ucl.ac.uk/WIBR/services/dna/). This service does not have a minimum sample number requirement and so was used for the final few reactions of this project. This service carries out sequencing reactions using a chain-termination mix from Beckman-Coulter. Sequencing PCRs were carried out by the service, so only the DNA template and primer had to be provided. Data was emailed to me as scf files. This did not affect the methods of analysis listed below.

3.2.7 Sequence data analysis

I used two different methods for analysing the sequence data. Firstly, for the sequencing of the *Cb-daf-2* and *Bm-daf-2* genes the abi data files for each gene were loaded into the SeqMan module of the Lasergene suite V5.51 (DNASTAR). This program is a contig generator and so as sequencing reads were added to the master file for each gene, the program would automatically align the reads to generate the most likely contiguous sequence. The final contigs represented the cDNA sequence of each gene. The contig was then entered into the EditSeq module of the Lasergene suite and translated *in silico* to generate the predicted protein sequence for these genes.

For the sequence analysis of the *daf-2* alleles I converted the abi data files to EditSeq files. This conversion allowed me to enter the sequence files that corresponded

to the + and – strand reads for each cDNA fragment into the MegAlign module of the Lasergene suite. This module aligned the reads for each fragment with the wild-type sequence of that fragment using the ClustalX algorithm and highlighted nucleotides in the reads that did not match the wild-type sequence. If a nucleotide did not match the wild-type sequence and was the same in both the + and – strand read it was treated as the putative mutation in that allele. The changed nucleotide was then incorporated into an EditSeq file with the rest of the *daf-2* cDNA wild-type sequence. This mutated cDNA was then translated *in silico* to generate the mutant protein sequence. The mutant protein sequence was then aligned with the wild-type DAF-2 sequence using the ClustalW algorithm and the program highlighted the changed amino acid.

3.2.8 Protein extraction

Protein samples for use in western blotting were prepared using a simple whole animal lysate method (Johnstone, 1999). Worms were grown to high-population density on a single 60mm plate and then washed into a 15 ml centrifuge tube with 3 ml of M9. The worms were then pelleted by centrifugation at 1000 rpm for 1 min. The supernatant was then removed and 5 volumes of SDS-sample buffer (Laemmli, 1970) were added to the estimated volume of worms (typically ~100 µl). The mixture was then transferred into a 1.5 ml centrifuge tube and boiled for 5 min at 100°C. Samples were then centrifuged at 14000 rpm in a refrigerated centrifuge at 4°C for 1 min to pellet the non-solubilised worms. Samples were then ready to run in a polyacrylamide gel.

3.2.9 Protein gels

A Mini-Protean 3 (Bio-rad) cell and casting assembly were used for SDS-polyacrylamide gel electrophoresis and casting, respectively, following the manufacturer's instructions. A range of different strength resolving gels were tried in the westerns and the mixes for each of these is listed in Table 3.2.2, whilst the stacking gel was always 3.75%.

The resolving gel was made up in a 15 ml centrifuge tube by adding the acrylamide solution, followed by the 3M Tris-HCl (pH 8.8) buffer, the 10% SDS

solution and finally the water. The solution was then gently mixed by inverting the tube several times. The casting cassette for the gel was prepared and once ready to receive the gel the ammonium persulphate and TEMED were added to the resolving gel mix. Again the solution was mixed by inverting the tube and then the solution was poured into the casting cassette with a disposable 10 ml pipette, up to the point where the top of the gel was 5mm from the bottom of the gel combs. The rest of the gel solution was discarded and the resolving gel in the casting cassette was overlaid with 1 ml of t-amyl alcohol, as recommended by the manufacturer.

The resolving gel was allowed to set for 1 hour before the t-amyl alcohol was drained off and the stacking gel poured on top. The stacking gel was made up in a 15 ml centrifuge tube and consisted of 280 µl of 40% acrylamide solution, 750 µl 0.5M Tris-HCl (pH 6.8), 30 µl 10% SDS solution and 1.914 ml water. As with the resolving gel, the stacking gel solution was gently mixed by inversion of the tube, before the addition of 22.5 µl 10% ammonium persulphate solution and 3 µl TEMED. Again the tube was mixed by inversion and the stacking gel was poured on top of the resolving gel with a 5 ml disposable pipette until it reached the top of the casting cassette. The stacking gel as then allowed to for 1hour before the gel combs were removed and the gel was loaded into the electrophoresis cell and filled with 1x Tris-glycine buffer (Laemmli, 1970). 20 µl of the relevant protein sample was then loaded in to each well in the stacking gel using gel-loading tips (ABgene). I used 10 µl of SeeBlue™ Plus 2 (Invitrogen) as my molecular weight ladder, which was always loaded into the first well on the left of the gel. The gel was then run at 100V until the bromophenol blue marker in the protein samples entered the resolving gel, at which point the voltage was increased to 120V. The gel was run until the bromophenol blue marker reached the end of the resolving gel. At this point the gel was ready for electrophoretic transfer of the proteins on to a blotting membrane.

Table 3.2.2 Resolving gel mixes

Components for 10 ml	5% Resolving Gel	7% Resolving Gel	10% Resolving Gel
40% Acrylamide (37:1)*	1.25 ml	1.75 ml	2.5 ml
3M Tris-HCl (pH 8.8)	1.25 ml	1.25 ml	1.25 ml
10% SDS (W/V)[a]	0.1 ml	0.1 ml	0.1 ml
Water	7.315 ml	6.815 ml	6.065 ml
10 % Ammonium persulphate (W/V)[a]	75 µl	75 µl	75 µl
TEMED	10 µl	10 µl	10 µl

* Ratio of acrylamide to bisacrylamide

[a] solutions made up with water

3.2.10 Western blotting

The protein gel was removed the electrophoresis cell. The stacking gel was discarded and the resolving gel placed into Towbin buffer to incubate for 30 min. Towbin buffer is made up of 1x Tris-glycine buffer (Laemmli, 1970) supplemented with methanol to a final concentration of 20% (V/V). A piece of Hybond nitrocellulose membrane (Amersham Biosciences), cut to be slightly larger than the gel, was incubated in Towbin buffer for 10 min. Two pieces of extra-thick Protean blotting paper (Bio-rad), cut to be slightly larger than the membrane, were incubated in Towbin buffer for 5 min. One piece of blotting paper was placed on the anode plate of a Trans-Blot Semi-Dry electrophoretic transfer cell (Bio-rad). The nitrocellulose membrane was then placed on top of the blotting paper. Air bubbles between the paper and the membrane were removed using a 5 ml disposable pipette. The gel was then placed on the membrane, so that the ladder remained on the left of the blot, again bubble were removed with a 5 ml pipette. The second piece of blotting paper was then placed on top of the gel. Bubbles were removed as before, then the cathode plate was put on top of the stack followed by the safety cover. The cell was run at 20V for 45 min

After removal from the Trans-Blot cell the gel was stained with Coomassie brilliant blue solution (Laemmli, 1970) to observe the efficiency of protein transfer. The membrane was stained with a 0.1% solution of Ponceau S dye in 5% acetic acid to confirm the presence of protein. The membrane was then destained by incubation in distilled water for 2 min. The membrane was then blocked overnight at 4°C with a 5% milk powder solution made up with PBS-Tween buffer.

On the following day the membrane was transferred to fresh 5% milk powder solution and incubated with the primary antibody at the required dilution for 90 min on a shaker at room temperature. After the primary antibody incubation the membrane was washed with a 15-minute incubation in PBS-Tween buffer followed by two 5-minute incubations. The membrane was then transferred to fresh 5% milk powder solution and incubated with the peroxidase-linked secondary antibody at the required dilution for 1 hour on a shaker at room temperature. After this incubation the membrane was washed with a 15-minute incubation in PBS-Tween buffer followed by four 5-minute incubations. Binding of the antibody was detected using the ECL plus chemiluminescence detection kit (Amersham Biosciences) according to the manufacturer's instructions. The ECL-treated membrane was placed into a BioMax cassette (Kodak) in a dark room. A sheet of Hyperfilm (Amersham Biosciences) was then exposed to the membrane in side the BioMax cassette for 1 min before being developed with a Kodak X-Omat developer. Depending on the outcome of the first exposure, a second piece of Hyperfilm was exposed for either shorter or longer periods of time to get the best exposure.

3.2.11 RNAi by ingestion of dsRNA

Two different protocols for RNA-mediated interference (RNAi) by ingestion of dsRNA, produced by transformed HT115(DE3) cells, were used in this project. The first was based on the 1mM IPTG-based induction method described by Kamath *et al* (2000). However, I supplemented the NGM plates with 25 μg/ml of carbenicillin instead of with 50 μg/ml ampicillin. Carbenicillin is an analogue of ampicillin that is more resistant to breakdown at room temperature.

The second method was lactose-based induction method, originally developed by Eric J. Lambie (personal communication) and optimised in our lab. NGM agar was prepared as described by Brenner (1974), except peptone concentration was reduced to a 10^{th} of standard from 2.5 g/L to 0.25 g/L. The agar was then supplemented with 0.2% lactose and 25 μg/ml carbenicillin. The peptone was reduced to prevent overgrowth of the HT115(DE3) cells as this strain of *E. coli* is not a uracil auxotroph like OP50, and is able to metabolise lactose.

Animals were grown on the RNAi treatment for at least three generations before phenotyping. All strains were maintained at 15°C before being shipped to the experimental temperature. This was done to prevent dauer formation in class 1 alleles of *daf-2*, which have an increased Daf-c phenotype, even on L4440 control plates, at 20°C (see chapter 6).

Chapter 4

Identifying *daf-2* homologues in the nematodes *Caenorhabditis briggsae* and *Brugia malayi*

4.1 Introduction

In this chapter I perform a comparative analysis of the *daf-2* gene, and the protein it encodes, with its homologues in three different nematode species. At the start of my project no other nematode *daf-2* sequences were available, which limits the understanding not only of the gene but also of features of the DAF-2 protein, such as its novel N- and C-terminal extensions. To gain further insight into *daf-2* function in *C. elegans* and its evolution within the nematode phylum, and to improve the protein sequence alignments used in modelling the structure of the DAF-2 receptor (see Chapter 5) I sequenced the homologues of this gene in *Caenorhabditis briggsae* and *Brugia malayi*, whose genomes were in the process of being sequenced. I have also created a model of the *daf-2* gene homologue, and its predicted protein, in a third nematode species *Caenorhabditis remanei*, whose genome is currently being sequenced.

4.1.1 The *daf-2* gene in *C. elegans*

The *daf-2* gene has 17 exons and spans 33.038 kb (Figure 4.2), the mRNA is 5795 bp and the translated protein is 1846 amino acids long. The predicted amino acid sequence shares 35%, 34% and 33% identity with the human insulin, IGF-1 and insulin-related receptors, respectively (Kimura *et al.*, 1997). The DAF-2 protein has a novel 121 amino acid (aa) N-terminal extension before the predicted signal peptide, an N-terminal extension is also found in dInR (Fernandez *et al.*, 1995; Ruan *et al.*, 1995), but not in any of the other characterised IRRs where the start codon is immediately followed by the signal peptide (Figure 4.3).

The DAF-2 receptor also has a large C-terminal extension (326aa following the kinase domain). Similar extensions are found in all characterised invertebrate IRRs, apart from the insulin receptor-like RTKs from *Sycon raphanus* (Skorokhod *et al.*, 1999). Most are smaller than DAF-2's, apart from the extension in dInR, which is ~400 amino acids long (Fernandez *et al.*, 1995; Ruan *et al.*, 1995). Another exception to this is the trematode *Schistosoma mansoni* SmRTK2 protein (AF314754). However, my own alignment of this protein with other IRRs suggests that the C-terminal sequence of this protein may be longer than reported.

A characteristic feature of these C-terminal extensions is the presence of multiple tyrosine residues, which may be phosphorylated upon receptor activation and act as binding sites for proteins with SH2 domains. For example, the C-terminal extension in DAF-2 contains 11 tyrosine residues, at least four of which may be putative SH2-binding sites (Kimura *et al.*, 1997). The C-terminal extension is thought to function similarly to the insulin receptor substrate (IRS) proteins, even though *C. elegans* does have a gene encoding an IRS homologue known as *ist-1* (Wolkow *et al.*, 2002).

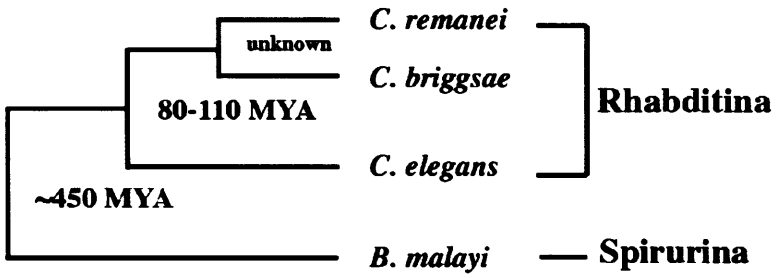
4.1.2 Identifying *daf-2* homologues in other nematodes

Of the seven invertebrate IRRs that had been sequenced prior to 2001, DAF-2 was the only sequence from the phylum nematoda. However, by this time two other nematode species were in the process of having their genomes sequenced. These were *Caenorhabditis briggsae* (Stein *et al.*, 2003) and *Brugia malayi* (Ghedin *et al.*, 2004). *Caenorhabditis briggsae* is also a soil dwelling nematode similar to *C. elegans* and has had its genome sequenced to act as tool for comparative genomics with *C. elegans*. The two species are thought to have diverged approximately 80-110 million years ago (MYA) during the Cretaceous period, but remain almost indistinguishable morphologically (Stein *et al.*, 2003). Both species are hermaphroditic, and like *C. elegans* low-frequency male formation also occurs in *C. briggsae*. Surprisingly though, this appears to be a convergent trait as the sister species of *C. briggsae*, *Caenorhabditis remanei* and the as yet unnamed *Caenorhabditis* species CB5161 are both dioecious, and all three species share a more recent ancestor than with *C. elegans* (Kiontke *et al.*, 2004). The genome of *C. remanei* is currently being sequenced and CB5161 is being prepared for sequencing.

Brugia malayi is a parasitic nematode that infects humans and causes lymphatic filariasis, commonly referred to as elephantiasis. This disease can be caused by three different parasitic nematodes. *Wuchereria bancrofti* is the major cause of this disease and is endemic in Africa, India, the Pacific islands, Caribbean, South America and Southeast Asia. *Brugia malayi* and the closely related *Brugia timori* are found only in Southeast Asia, *B. timori* being restricted to eastern Indonesia (Cox, 2000). Even though *W. bancrofti* is the major cause of lymphatic filariasis, *B. malayi* was chosen for genome sequencing due to fact that it can be cultured in laboratory conditions (Ghedin

et al., 2004). In terms of *B. malayi* relatedness to *C. elegans*, the two species are estimated to have diverged approximately 450 MYA, during the Ordovician period, based on sequence divergence of their globins (Vanfleteren *et al.*, 1994); M. L. Blaxter, personal communication). However, the divergence dates for both *C. briggsae* and *B. malayi* from *C. elegans* should be treated with some scepticism as there is no fossil record for the nematode phylum with which to calibrate the molecular clock. In the context of IIS and the study of ageing it is notable that the lifespan of *B. malayi* adults is at least 6-7 years (Gems, 2000).

Figure 4.1 Phylogenetic relationship between *Brugia* and *Caenorhabditis*



A simplified phylogenetic tree that shows the relationship between *Brugia malayi* (clade Spirurina) and *Caenorhabditis elegans*, *Caenorhabditis briggsae* and *Caenorhabditis remanei* (clade Rhabditina). *B. malayi* is thought to have diverged from the *Caenorhabditis* species ~450 million years ago (MYA). *C. briggsae* and *C. remanei* are sister species whose phylogenetic relationship with each other is currently unresolved (Cho *et al.*, 2004; Kiontke *et al.*, 2004), they are thought to have diverged from *C. elegans* 80-110 MYA (Stein *et al.*, 2003).

4.1.3 The value of comparative analysis of nematode insulin/IGF-1R receptors

The existence of these three nematode genome sequencing projects provides an opportunity for better understanding of nematode insulin/IGF-1R receptors, since it allows rapid identification of the *daf-2* orthologues in each species. By comparing the amino acid sequences of the DAF-2 protein from the three nematode species with other IRRs I hoped to identify regions that have a functional role in all IRRs, the nematode phylum or only in the *Caenorhabditis* genus. The ability to compare the genomic

sequence around the *daf-2* locus in multiple species of nematode may also help to identify conserved regulatory elements, especially as the *daf-2* promoter in *C. elegans* remains uncharacterised.

4.2 Materials and Methods

4.2.1 Strains and cDNA samples

The following nematode strains were used for sequencing the *C. briggsae* and *B. malayi daf-2* homologues: AF16 *Caenorhabditis briggsae* (Gujarat) (Fodor *et al.*, 1983) obtained from CGC and *Brugia malayi* TRS strain (Foster *et al.*, 2004), two independent samples were provided by both Yvonne Harcus & Rick Maizels, University of Edinburgh and David Guiliano & Murray Selkirk, Imperial College. An aliquot of first strand cDNA (kindly provided by Mark L. Blaxter, University of Edinburgh) and a λ phage cDNA library of *B. malayi* microfilariae (kindly provided by Murray Selkirk) were also used as template in the sequencing of the *B. malayi daf-2* homologue.

4.2.2 Identification of genomic region for *daf-2* homologues

Candidate genomic regions for the *daf-2* homologues in both *C. briggsae* and *B. malayi* were identified by performing a TBLASTN (Altschul *et al.*, 1990) search of the respective genome assemblies using the amino acid sequence of the *C. elegans* DAF-2 protein. The search for the *C. briggsae daf-2* homologue was carried out on the cb25.AGP8 assembly of the *C. briggsae* genome (http://www.sanger.ac.uk/cgi-bin/blast/submitblast/c_briggsae). The *B. malayi* search (http://tigrblast.tigr.org/er-blast/index.cgi?project=bma1) was carried out on the BMA1 scaffold assembly of the genome. The TBLASTN search for the *C. remanei daf-2* homologue was carried out using the pCAP v2 assembly of the genome (http://www.genome.wustl.edu/projects/cremanei/cremanei_BLAST/cremanei_client.pl) with both the DAF-2 and *Cb*-DAF-2 (first version, see 4.31 for details) protein sequences. The DNA sequence of the top BLAST hit in each case was then analysed *in silico* to model the structure of the homologous *daf-2* gene (see below).

4.2.3 Modelling of the gene structure of *daf-2* homologues

The high scoring pairs (HSP) within each of the best BLAST hits were imported as individual protein sequence files into the EditSeq™ module of the Lasergene suite v5.51 (DNASTAR). These HSP protein files were then aligned against their respective contig using the protein finder function in the GeneQuest™ module of Lasergene and treated as putative exons. This procedure resulted in a preliminary model of the gene structure for all three *daf-2* homologues. The consensus splice donor and acceptor sites as determined for *C. elegans* were then inputted into the contig sequences. Putative cDNA sequences for each of the *daf-2* homologues were constructed by linking splice donor and acceptor sites that were in frame with the aligned HSP/ putative exon segments.

4.2.4 Initial primer design

Initial primers for the amplification of cDNA fragments from the *C. briggsae* and *B. malayi daf-2* transcripts were designed based on the putative cDNA models constructed in GeneQuest™. Primers were designed using the Primer3 WWW interface (Rozen and Skaletsky, 2000) (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Primers were generally designed to amplify overlapping fragments 350 bp-500 bp in size. A list of the main primer pairs used in the sequencing of the *C. briggsae* and *B. malayi daf-2* cDNAs can be found in Table 4.2.1 and 4.2.2, respectively.

4.2.5 Sequencing of the *Cb-daf-2* and *Bm-daf-2* genes

Total RNA from *C. briggsae* mixed stage worms was isolated using the same procedure as that employed for *C. elegans* (see chapter 3.2.1). Total RNA from adult *B. malayi* was extracted by placing 3-12 worms into a Lysing Matrix D tube (Q-biogene) containing 1ml of TRIzol (Invitrogen). The worms were then subjected to two 30-second cycles of homogenisation in FP220 ribolyser (Q-biogene) using a speed setting of 4.5. Subsequent extraction of RNA from the TRIzol homogenate was carried out according to the manufacturers instructions. Methods for first strand cDNA synthesis from the RNA and amplification of and sequencing of overlapping fragments by PCR

can be found in chapter 3 (sections 3.2.2, 3.2.4 and 3.2.6). The sequencing reads were assembled into contiguous cDNA sequences using the method described in chapter 3.2.7.

The confirmed cDNA sequences for the two *daf-2* genes were then aligned against the genomic contigs, from the BLAST searches, using the MegAlign™ module of the Lasergene suite. Alignments were performed with a ClustalX algorithm. The alignments were then hand adjusted to take account of consensus splice sites. This process then gave the final genomic structure of the gene.

4.2.6 Protein sequence alignments

All protein alignments were carried using the default settings on the ClustalW (Thompson *et al.*, 1994) server (http://www.ebi.ac.uk/clustalw/). The alignments were then hand-adjusted using the BioEdit© sequence-editing programme (Tom Hall, http://www.mbio.ncsu.edu/BioEdit/bioedit.html).

Table 4.2.1 Primer pairs used in sequencing *Cb-daf-2* cDNA:

5' primer	5'-sequence-3'	3' primer	5'-sequence-3'	Product size
SL1	GGTTTAATTACCCAAGTTTGAG	F1sub1R	CTACGTCGCGTTCTGTGACTAC	440 bp
F1sub1F	TGAGCACAATCACCTCCTCTAA	F1R	AGTGCCAGTGACTTCTCGCAAA	582 bp
F2F	CATTCATCTATGGCGGGGGACA	F2R	AGTCTCCGTCACCGACCACTCG	395 bp
F3F	TGTGGTTCGAAACGGTGGCATC	F3R	GAGCATCACAGCCACCCAGACA	362 bp
F4F	CGCCAGAAGGAGAACGATGTCA	F4R	CCAACGCCTTGGGAAATGTGTC	363 bp
F5F	AAATGCGATGGACCGTGTGAGA	F5R	TGGCATTCTTGTTGCCAAACACA	367 bp
F6F	GCGATCGCTTTTTGAGGATAATCG	F6R	TGAACACCGATGACCACGAGTC	386 bp
F6F*	GCGATCGCTTTTTGAGGATAATCG	F7R	GCGTGCACATCGATATCCAGCTCT	693 bp
F8F	TGGGATCCACCACACCATAGCA	Cbrig5'rev	TCCTGGTCCTTTTCGTTGGGTGT	431 bp
F9F	GTCGCCACCCGTTATGAAGACG	F10R	TGGCATGCTTGGATTGTGATCG	569 bp
F11F	TGGTTGTTGGGGAATCTGAAGCA	F11R	TTCGGACAAATCGCGGAATACG	354 bp
F12F	TTCGATCCAACCACCCATGGTA	F12R	CCTCTCCCAGTGTCACCGACTCC	355 bp
F13F	CCATTGACGATTGGGAGCTGGA	F13R	TCGTCGAATACCTCCTCCTCTCG	352 bp
F14F	TGAGAGCCAAGCGAGAGGAGGA	F14R	GCCGAGGGTGATCAGCTCGTAG	368 bp
F15F	CGACATTTGGAGTTTCGGAGTGG	F15R	GTCCAGCTCCTCAGGCGGATTA	352 bp
F16F	CCATGGAGGACACGGAACCACT	F16R	TAGGCACATCTGCTGCCACGTC	384 bp
F17F	CACAGTCGTGGACCGAGTGACA	F17R	CCGCCTGCATTTGCTTCTGTTT	394 bp
F18F	GAGAGGGACTACGGTTGACAGA	Exon26R	CTCCTCGATGAGACGTTCTGTT	443 bp
Exon25-26fwd	AAGGATCGCCTTCGAAAAAT	Oligo dT	TTTTTTTTTTTTTTTTTVN	Incomplete

Primer pairs used to sequence *Cb-daf-2* cDNA sequence. Fragments are overlapping and are ordered from the 5' to the 3' end of the sequence. *The F7F primer was designed based on the model of the gene structure, however this primer sequence did not anneal to the actual cDNA, so F6F was used with F7R instead.

Table 4.2.2 Primer pairs used in sequencing *Bm-daf-2* cDNA:

5' primer	5'-sequence-3'	3' primer	5'-sequence-3'	Product size
SL1	GGTTTAATTACCCAAGTTTGAG	BM1R	TCAACAGATGTCCTGTAATGACG	359 bp
BM1F	TCAAGACAAGGAAACGAATGAA	BM11R	GCTGGTCCGATGAGTATTTGA	390 bp
Exon31longfwd	TGAGTGCTCTTCGATCGTTAAA	BM4R	CCTTGGTAAACGACATTTTTGC	511 bp
BM4F	GACCAGGTTGTGACGATGATG	BM2R	TTCCTTTTTCTGGTTCGAAATG	1038 bp
BM21F	CCGATATGGATCATCGAAAA	BM5R	ACCCTGTATGCAAGTATGATGC	1147 bp
BMmidFwd	TGAAAATGCCGTTCAAAATCTA	Exon20rev	TGCACATTCACCGAAGGTAA	1030 bp
Exon17fwd	GGAACGAGAAGGTGTGGTGT	BM6R	ATACCAAAAATCCGGACAACC	1096 bp
BM3F	CTACGTGACTACCTGCGGTCTA	BMendrev	TGTCCTGCACTAGCAGTTCGTA	707 bp
Ctermfwd	TCCAATCAAATGAATACGGATG	Oligo dT	TTTTTTTTTTTTTTTTTTVN	525 bp

Primer pairs used to sequence *Bm-daf-2* cDNA sequence. Fragments are overlapping and are ordered from the 5' to the 3' end of the sequence.

4.3 Results

At the start of my project the only nematode gene encoding a member of the IR superfamily of proteins was *daf-2* in *C. elegans*. The DAF-2 protein has novel extensions at both termini, whose function is unknown. To gain further insight into the novel features of the DAF-2 protein and the evolution of its gene, I sequenced its orthologues in the nematodes *C. briggsae* and *B. malayi*. I have also modelled the gene structure of the *daf-2* orthologue in *C. remanei*, whose genome sequence is currently being assembled.

4.3.1 Identification of the *Caenorhabditis briggsae daf-2* homologue

The TBLASTN search of the *C. briggsae* genome with the DAF-2 amino acid sequence yielded the following ~1.4Mb supercontig cb25.FPC4030 (EMBL: CAAC01000074) with an E-value of $8.6e^{-282}$. This contig was analyzed *in silico* to create a model of both the gene and its cDNA (see methods). From the BLAST search there were no HSP alignments for amino acids N-terminal of residue 135 or C-terminal of residue 1805 of the *C. elegans* DAF-2 sequence. Therefore the cDNA model for *C. briggsae daf-2* sequence was incomplete at the 5' and 3' ends and did not have putative start and stop codons or either of the untranslated regions (UTRs).

To help amplify the 5' end of the cDNA sequence I used a primer with the sequence of the SL1 trans-spliced leader in conjunction with a cDNA specific primer. The SL1 trans-spliced leader is a short 22-nucleotide RNA segment that is spliced onto the 5' end of many mRNAs, not only in *C. elegans*, but also through out the nematode phylum (Blumenthal and Steward, 1997). To amplify the 3' end of the cDNA sequence I used an oligo-δT primer in conjunction with a cDNA-specific primer.

Two PCR reactions were performed per cDNA fragment, and both strands of each amplicon were sequenced, giving 4-fold coverage of the cDNA sequence of each fragment. As the cDNA fragments were designed to overlap, many parts of the cDNA sequence were actually covered by eight different reads. The sequencing reads confirmed that there is an additional thymidine nucleotide between positions 598450

and 598451 of the cb25.FPC4030 contig. This additional nucleotide is essential in preserving the frame of the *C. briggsae daf-2* cDNA sequence.

The original *C. briggsae daf-2* (*Cb-daf-2*) sequence (GenBank: AY382557) was deposited in GenBank database on 05/09/2003. However, this sequence will be updated after submission of this thesis due to subsequent identification of an additional exon at the 3' end of the gene. This resulted from a TBLASTN search of the *C. briggsae* genome using the *Cr*-DAF-2 sequence (see below), which suggested that the submitted *Cb-daf-2* gene might have an additional exon at the 3' end. This additional exon has been confirmed by sequencing; however, the sequence of 3' UTR was still incomplete at the time of writing this thesis. The *Cb-daf-2* gene is located on the minus strand of the cb25. FPC4030 supercontig. The *Cb-daf-2* gene is greater than 27 kb in length with 26 exons (Table 4.3.1 and Figure 4.2). The cDNA is 5618 bp in length, excluding the 3' UTR, and contains an open reading frame (ORF) of 5424 bp, which encodes a protein of 1807 amino acids (see Figure 4.3).

4.3.2 Identification of the *Brugia malayi daf-2* homologue

The *B. malayi* search of the BMA1 scaffold assembly using the DAF-2 amino acid sequence returned the following ~34 kb contig TIGR_assembly_14419 (CA: 1384396) with an E-value $1.5e^{-118}$. For *B. malayi* there were no HSP alignments N-terminal of residue 146 and C-terminal of residue 1734 of the DAF-2 sequence. So like the *Cb-daf-2* cDNA model, this model was also incomplete at the 5' and 3' ends. The same sequencing strategy that was used to obtain the ends of the *Cb-daf-2* cDNA sequence was employed for the *B. malayi daf-2* homologue. Again, the majority of the cDNA sequence had 4-fold coverage rising to more than 10-fold in certain parts of the cDNA. It should be noted that I was greatly assisted in the sequencing of the *B. malayi daf-2* homologue by Nachiket Nadkarni, a former summer student in the Gems lab under my supervision.

Our sequencing of the *B. malayi daf-2* homologue identified a number of sequencing errors in the TIGR contig. The sequencing reads identified the anonymous nucleotides at positions 6147 and 7108 of the 14419 contig as cytidines and another at position 8453 was a thymidine. There were two single nucleotide insertions in the contig: an additional guanosine between positions 7359 and 7360 and an additional

adenosine between positions 12127 and 12128. There is also strong evidence that the adenosine nucleotide at position 7194 of the 14419 contig should be deleted from the sequence. This is because this nucleotide is not present in the cDNA sequence of the *Bm-daf-2* gene and disrupts the consensus splice donor site for exon 12 of the gene.

The *B. malayi daf-2* (*Bm-daf-2*) (GenBank: AY856075) gene is located on the plus strand of the TIGR_assembly_14419 contig, starting at position 2667 and ending at 13868. The *Bm-daf-2* gene is 11202 bp and spans 30 exons (see Table 4.3.2 and Figure 4.2). The cDNA is 4617 bp long with an ORF of 4371 bp that translates into protein of 1456 amino acids (see Figure 4.3).

4.3.3 Identification of the *Caenorhabditis remanei daf-2* homologue

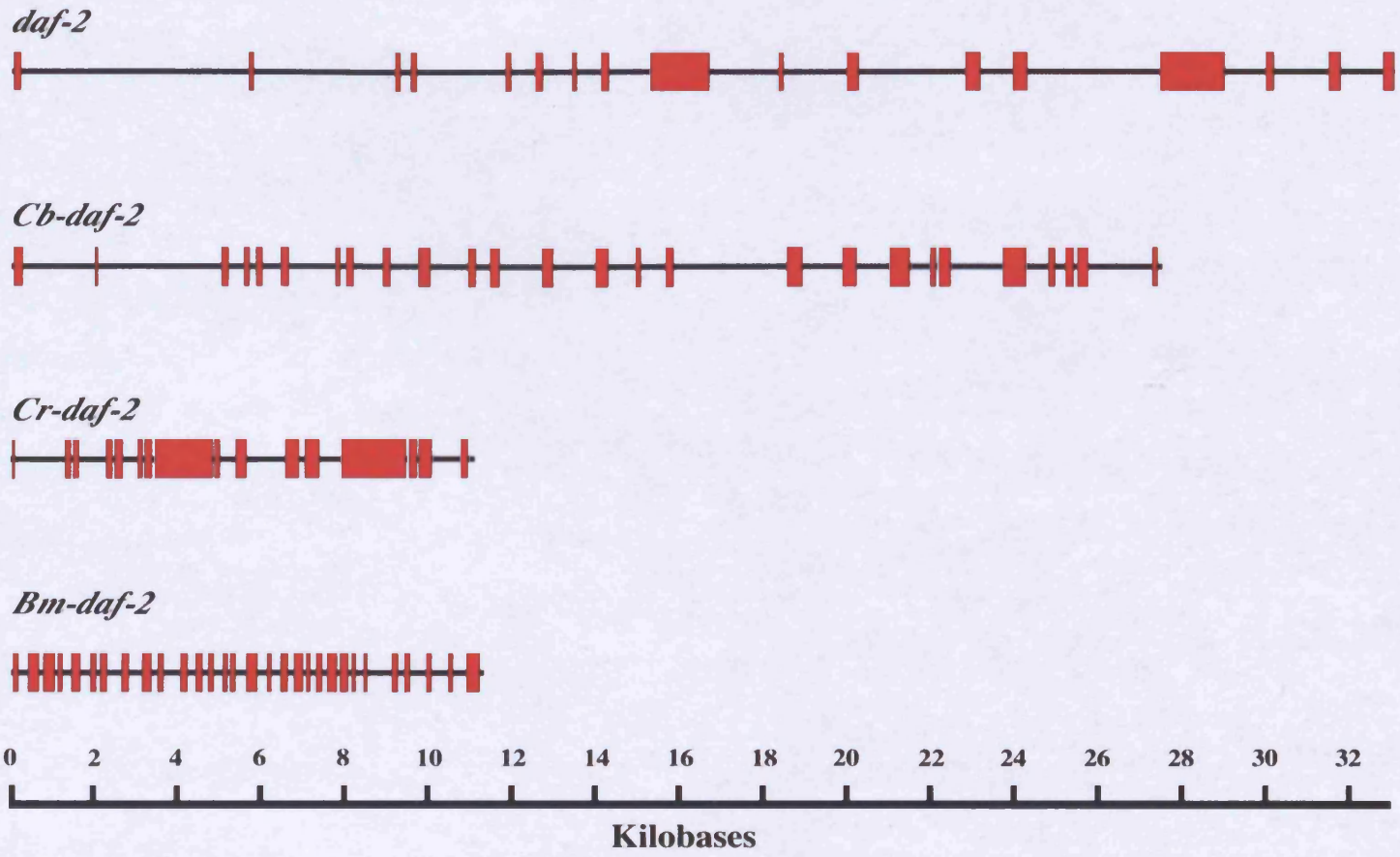
Recently the genome of *Caenorhabditis remanei* has also been sequenced and is in the process of being assembled into large contigs. Late in this project I performed a TBLASTN search of the pCAP v2 assembly of the *C. remanei* genome using both the *Ce*-DAF-2 and *Cb*-DAF-2 (first version) sequences identified a ~36 kb contig named 5.28 with an E-value of 0. Using the same methodology as described for the modelling of the *Cb-daf-2* and *Bm-daf-2* genes it was possible to develop a model of the putative *C. remanei daf-2* (*Cr-daf-2*) gene.

The *Cr-daf-2* gene model includes a putative start and stop codon, unlike the models of the *Cb-daf-2* and *Bm-daf-2* genes, but does not include the 5' and 3' UTRs. The putative terminal exon and putative stop codon of the *Cr-daf-2* gene were based on a HSP alignment of the C-terminus of the *Ce*-DAF-2 sequence, which represents the terminal exon and stop codon of the *Ce-daf-2* gene. In order to model the putative start codon the 11215 bp upstream of the predicted second exon of the *Cr-daf-2* gene was scanned for all DNA sequences that contained a start codon and splice donor site that were in frame with the second exon. Only one such sequence was present in the entire upstream sequence. This 59 bp segment was therefore taken to be the putative first exon, assuming that the 5' UTR of the *Cr-daf-2* gene is not interrupted by an intron.

Unfortunately, due to time-constraints this model could not be confirmed by sequencing the cDNA. However, based on this model the *Cr-daf-2* gene, excluding the 5' and 3' UTRs, is 11605 bp in length and has 16 exons (Table 4.3.3 and Figure 4.2), assuming that neither of the UTRs is interrupted by an intron. The *Cr-daf-2* gene is

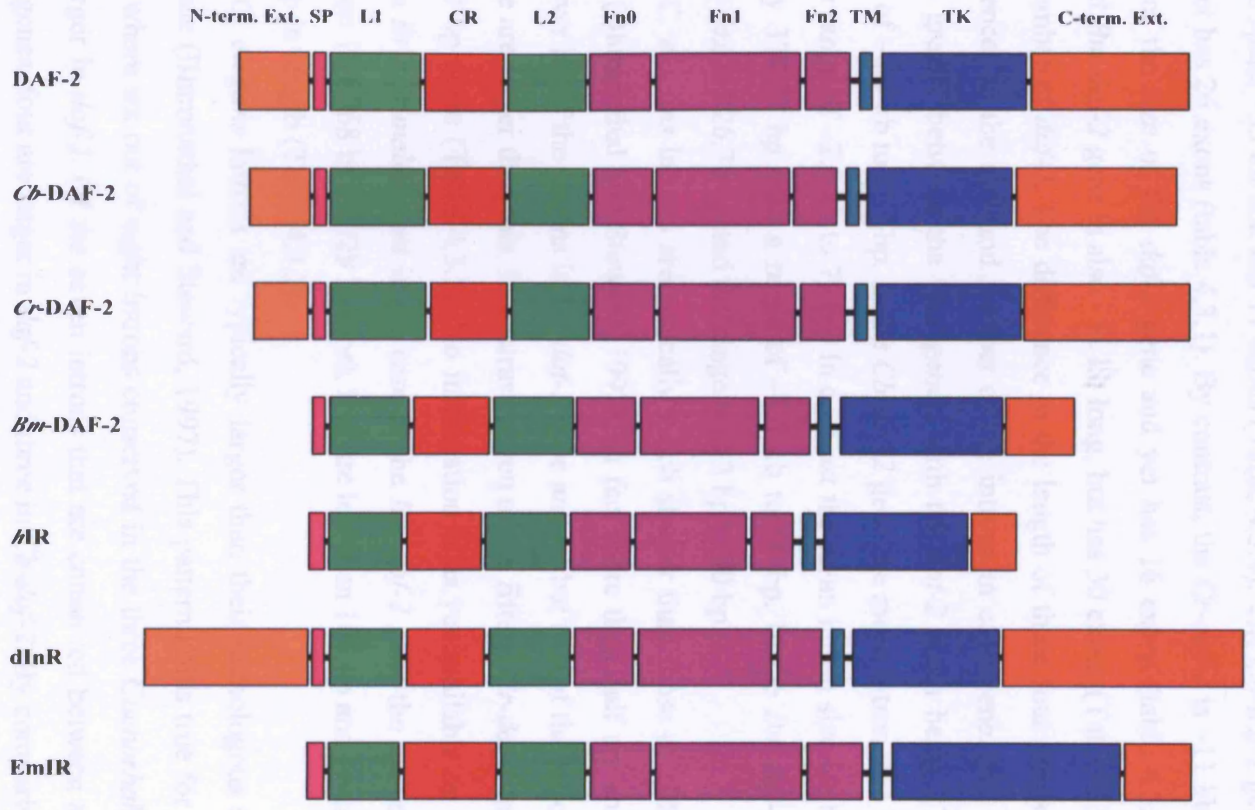
located on the minus strand of contig 5.28, the coding sequence of the first exon begins at position 26345 and the stop codon in the terminal exon ends at position 15281. The predicted *Cr*-DAF-2 protein is 1816 amino acids long (Figure 4.3).

Figure 4.2 Structure of the nematode *daf-2* genes



A schematic representation of the gene structure of the *daf-2* gene in four different species of nematode. The red rectangles represent the exons of each *daf-2* gene, while the interconnecting lines represent the introns. The *daf-2* gene from *C. elegans* is ~33 kb in length and has 17 exons. The *Cb-daf-2* gene from *C. briggsae* is ~27 kb in length and has 26 exons. The *Cr-daf-2* gene from *C. remanei* and the *Bm-daf-2* gene from *B. malayi* are both ~11 KB in length and have 16 and 30 exons, respectively.

Figure 4.3 Comparison of the nematode DAF-2 proreceptors



Comparison of the nematode DAF-2 proreceptors with the human insulin receptor (*h*IR), *D. melanogaster* insulin receptor (dInR) and the *E. multilocularis* insulin receptor (EmIR) proreceptors. The proteins are aligned by their signal peptides (SP) so that the differences in their architecture can readily be seen. All seven proteins have near identical domain architectures, except DAF-2, *Cb*-DAF-2, *Cr*-DAF-2 and dInR, which possess large N-terminal extensions before the signal peptide.

4.3.4 Comparison of the gene size of the four nematode *daf-2* genes

A comparison of the exon-intron structure of the four nematode *daf-2* genes revealed a striking and unexpected variation in the gene structure (Figure 4.2). The *C. elegans daf-2* gene spans ~33 kb and has 17 exons (Table 4.3.4). The *Cb-daf-2* gene is ~27 kb long and yet has 26 exons (table 4.3.1). By contrast, the *Cr-daf-2* is ~11 kb in length, only a third of the size of the *daf-2* gene and yet has 16 exons (table 4.3.3). The distantly related *Bm-daf-2* gene is also ~11 kb long, but has 30 exons (Table 4.3.2), almost twice the number of *daf-2*. The difference in the length of these four genes is mainly due to differences in the size and number of the introns in each gene. The mean intron size varies greatly between the four genes, with the *daf-2* mean being 1702.63 bp with a range of ~5.5 kb to 257 bp. In the *Cb-daf-2* gene the mean intron size is only 868.16 bp with a range of ~2.9 kb to 79 bp. In contrast the mean intron size for the *Cr-daf-2* gene is only 374.27 bp with a range of ~1.5 kb to 51 bp. In the *Bm-daf-2* gene the mean intron size is 226.72 bp and the range is 603 bp to 80 bp.

C. elegans introns are typically much shorter than those of either vertebrates or yeast (Blumenthal and Steward, 1997). In fact more than half are shorter than 60 bp. However all of the introns in the *daf-2* gene and all but two of the introns in the *Cb-daf-2* gene are larger than this. In contrast, seven of the fifteen *Cr-daf-2* introns are close to the 60 bp norm (Table 4.3.3). No information is as yet available on the modal intron size in *Brugia malayi*, but in the case of the *Bm-daf-2* gene the majority of introns are between 114-268 bp (17/29 introns), five are less than 100 bp and seven are greater than 300 bp in length (Table 4.3.2).

C. elegans introns are typically larger than their homologous counterpart in *C. briggsae* (Blumenthal and Steward, 1997). This pattern holds true for *daf-2* versus *Cb-daf-2* where six out of eight introns conserved in the three *Caenorhabditis daf-2* genes are larger in *daf-2*. Of the seven introns that are conserved between all four nematode *daf-2* genes, four are larger in *daf-2* and three in *Cb-daf-2*. By comparison, all of the *Cr-daf-2* introns that are conserved with the other *Caenorhabditis daf-2* genes are shorter than their *C. elegans* and *C. briggsae* counterparts. The *Bm-daf-2* gene has no conserved introns that are larger than those of *C. elegans*, but one intron larger than *C. briggsae* and three larger than *C. remanei*.

The major cause of the variation in the size of introns in the *daf-2* gene appears to the accumulation of repetitive DNA sequences within each intron. For example the first intron of *daf-2* is 5464 bp long, of which 2319 bp (42.4%) consists of repetitive elements identified using the RepeatMasker programme (http://www.repeatmasker.org) with the *C. elegans* repeat library. In total the amount of repetitive sequence in the intronic sequence of the *daf-2* gene is 13040 bp (47.9%). There are no species-specific libraries available yet for the other three nematodes, so it is not possible to do this analysis on their intronic sequence.

Table 4.3.1 Exon- Intron sizes of *Cb-daf-2* gene

Exon No.	Exon Size (bp)	Intron No.	Intron Size (bp)	Phase of Intron*	Amino Acid interrupted§
1	206	1	1741	0	W4: I5
2	57	2	2959	2	S24
3	162	3	378	0	M77: I78
4	120	4	166	0	Q117: P118
5	137	5	453	2	S163
6	183	6	1132	2	K224
7	109	7	150	0	A260: I261
8	169	8	715	1	V317
9	172	9	685	1	G374
10	265	10	907	0	K462: D463
11	174	11	339	0	K520: E521
12	219	12	1033	0	Q593: F594
13	246	13	1024	0	K675: E676
14	299	14	663	2	K775
15	116	15	603	1	G814
16	184	16	2710	2	R875
17	373	17	957	0	A999: I1000
18	326	18	814	2	E1108
19	472	19	496	0	K1265: T1266
20	120	20	79	0	E1305: V1306
21	273	21	1256	0	K1396: D1397
22	552	22	542	0	K1580: P1581
23	141	24	262	0	T1627: R1628
24	186	24	92	0	T1689: R1690
25	246	25	1548	0	S1771: K1772
26	111	Total Intronic sequence: 21704 bp			

*The phase of intron refers to the codon position interrupted by an intron: phase 0 means the intron lies between the third base of one codon and the first of the next codon, phase 1 introns interrupt the codon after the first base and phase 2 introns interrupt after the second base of the codon.

§Using single letter amino acid code and treating the start codon as the first amino acid position

Table 4.3.2 Exon- Intron sizes of *Bm-daf-2* gene

Exon No.	Exon Size (bp)	Intron No.	Intron Size (bp)	Phase of Intron	Amino Acid interrupted
1	117	1	255	0	S13: L14
2	233	2	124	2	S91
3	263	3	80	1	D177
4	95	4	225	0	L210: E211
5	208	5	249	2	R280
6	125	6	114	1	D322
7	148	7	360	2	R371
8	162	8	329	2	R425
9	209	9	170	1	C495
10	117	10	423	0	Q533: V534
11	164	11	205	2	W588
12	152	12	146	1	V639
13	120	13	229	1	G679
14	105	14	91	0	K775
15	101	15	268	2	Q747
16	250	16	258	0	W830: I831
17	88	17	236	1	A860
18	161	18	158	0	Q913: T914
19	198	19	85	0	T979: V980
20	93	20	143	0	K1010: V1011
21	141	21	141	0	K1057: V1058
22	222	22	92	0	K1131: G1132
23	177	23	95	0	M1190: V1191
24	82	24	204	1	G1218
25	79	25	603	2	W1244
26	138	26	160	2	W1290
27	130	27	393	0	E1333: M1334
28	123	28	408	0	E1374: T1375
29	105	29	331	0	K1409: P1410
30	312	Total Intronic sequence: 6575 bp			

Table 4.3.3 Exon- Intron sizes of *Cr-daf-2* gene

Exon No.	Exon Size (bp)	Intron No.	Intron Size (bp)	Phase of Intron	Amino Acid interrupted
1	59	1	1460	2	K20
2	127	2	68	0	M62: T63
3	126	3	657	0	Q104: P105
4	137	4	72	2	T150
5	189	5	357	2	T213
6	109	6	53	0	A249: I250
7	169	7	70	1	L306
8	1378	8	57	2	L765
9	116	9	383	1	A804
10	253	10	927	2	R888
11	322	11	153	0	G995: I996
12	335	12	529	2	E1107
13	1549	13	83	0	T1623: R1624
14	180	14	51	0	T1683: R1684
15	300	15	694	0	S1783: K1784
16	102	Total Intronic sequence: 5614 bp			

Table 4.3.4 Exon- Intron sizes of *daf-2* gene

Exon No.	Exon Size (bp)	Intron No.	Intron Size (bp)	Repetitive sequence of Intron (bp, %)	Phase of Intron	Amino Acid interrupted
1	162	1	5464	2319 (42.44%)	0	E21: E22
2	97	2	3404	1026 (30.14%)	2	R54
3	115	3	257	38 (14.79%)	0	T92: T93
4	132	4	2127	1259 (59.19%)	0	Q136: P137
5	122	5	578	440 (76.12%)	2	R177
6	177	6	697	29 (4.16%)	2	E236
7	106	7	605	28 (4.63%)	0	E271: I272
8	170	8	1009	639 (63.33%)	1	M328
9	1393	9	1662	1319 (79.36%)	2	H792
10	113	10	1518	1455 (95.85%)	1	A830
11	283	11	2572	1067 (41.49%)	2	R924
12	346	12	791	578 (73.07%)	0	A1039: I1040
13	335	13	3183	1567 (49.23%)	2	F1151
14	1531	14	1004	738 (73.51%)	0	T1661: R1662
15	186	15	1326	426 (32.13%)	0	T1723: R1724
16	276	16	1045	112 (10.72%)	0	S1815: R1816
17	252	Total Intronic sequence: 27242 bp		Total Repetitive Sequence: 13040 bp (47.87%)		

4.3.5 Exon- intron structure and evolution:

Recently it has become apparent that intron loss is a major feature of nematode evolution, especially within the *Caenorhabditis* genus (Cho *et al,* 2004; Kiontke *et al,* 2004). At first glance the exon-intron structure of the three *Caenorhabditis daf-2* genes compared to that of *Bm-daf-2* would agree with this phenomenon, given that *daf-2*, *Cb-daf-2* and *Cr-daf-2* have 16, 25, and 15 introns respectively, compared to the 29 introns found in *Bm-daf-2*. Similarly one might easily jump to the conclusion that there had been extensive intron loss in both *daf-2* and *Cr-daf-2* compared to *Cb-daf-2*. However, neither conclusion is correct. By mapping the position of the introns of the four *daf-2* genes onto a protein alignment including three other IR sequences (*h*IR, DInR and EmIR) to act as outgroups, one can use intron conservation to map losses and potentially intron gains (Table 4.3.5). Using this technique it becomes apparent that at least 5 introns, *Bm-daf-2* introns 4, 9, 17, 25 and 26, have been unambiguously lost from the *Caenorhabditis daf-2* genes when compared to the *Bm-daf-2* gene. Conversely only one intron has been lost from the *Bm-daf-2* gene when compared to the *Caenorhabditis daf-2* genes, equivalent to *daf-2* and *Cb-daf-2* intron 7 and *Cr-daf-2* intron 6. When compared to introns conserved between *h*IR, DInR and EmIR or *h*IR and EmIR there is evidence that all four nematode *daf-2* genes have lost five introns equivalent to introns 1, 4, 7, 8 and 15 from the *h*IR gene.

Although intron loss is a major feature of nematode evolution all three *Caenorhabditis daf-2* genes have also been subject to intron gain, especially *Cb-daf-2* when compared to *Bm-daf-2*. The first three introns of the *daf-2* and *Cb-daf-2* genes and the intron 1 and 2 from *Cr-daf-2* all map to the novel N-terminal extension that seems to be a feature of *Caenorhabditis* DAF-2 proteins. Because this feature is not present in the *Bm*-DAF-2 protein, it can be assumed that these introns must be *Caenorhabditis* specific gains. Similar intron gains can also be seen at the C-terminus of the *Caenorhabditis* DAF-2 proteins. The homologous introns 15, 16 in *daf-2* and 24, 25 in *Cb-daf-2* and 14, 15 in *Cr-daf-2* all map to the C-terminus of their respective proteins. These regions are not present in the C-terminus of the *Bm*-DAF-2 protein and therefore must also be *Caenorhabditis* specific gains.

When comparing the exon-intron structure of the three *Caenorhabditis daf-2* genes, the structure of the *Cb-daf-2* gene stands out as being very different from that of

daf-2 and *Cr-daf-2*. The *daf-2* and *Cr-daf-2* genes have a near identical gene structure, they only differ by the presence of an additional intron, intron 2, in the *daf-2* gene. At present it is impossible to ascertain whether this intron represents a gain in *daf-2* or a loss in *Cr-daf-2*. Whereas the *Cb-daf-2* gene has 10 additional introns compared to the other two *daf-2* genes, these are introns 1, 9, 10, 11, 12, 13, 19, 20, 21 and 22. Given that *C. briggsae* and *C. remanei* are likely sister-species that share a more recent ancestor than with *C. elegans* (Cho *et al.*, 2004; Kiontke *et al.*, 2004) and that the *Cr-daf-2* gene structure resembles that of *daf-2* more than that of *Cb-daf-2* it is quite likely that these additional introns all represent recent gains.

Table 4.3.5 Position of introns across aligned protein sequences

Intron position	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
DAF-2		1	2	3	4		5	6	7	8															
Cb-DAF-2	1	2		3	4		5	6	7	8		9				10					11		12		
Cr-DAF-2		1		2	3		2	5	6	7															
Bm-DAF-2					1			2		3	4			5	6				7			8		9	10
*h*IR						1					2						3			4		5		6	
DInR						1					2		3							4				5	
EmIR						1			2		3							4		5		6		7	

Intron position	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
DAF-2						9	10			11				12						13					
Cb-DAF-2	13					14	15			16				17						18					
Cr-DAF-2						8	9			10				11						12					
Bm-DAF-2				11		12	13		14	15				16		17		18						19	
*h*IR		7			8		9				10	11				12			13			14			15
DInR							7									8					9				
EmIR		8	9		10		11	12					13		14	15	16						17		18

Intron position	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
DAF-2																						14		15	16
Cb-DAF-2					19						20					21					22	23		24	25
Cr-DAF-2																						13		14	15
Bm-DAF-2	20		21						22				23		24		25	26	27	28			29		
*h*IR		16				17				18				19			20	21							
DInR																									
EmIR				19			20	21				22					23	24							

This table shows the positions (red numbers) at which introns are present in the aligned sequences of all 4 nematode DAF-2 proteins and 3 outgroup members of the IR superfamily (*h*IR, DInR and EmIR). The numbers in black are the actual introns from each gene that map to this position of the protein alignment. By mapping the position of each intron onto the protein alignment it is possible to calculate which introns have been gained and lost in the nematode *daf-2* genes.

4.3.6 Identifying regulatory elements in the *daf-2* genomic sequence

One of the major reasons that related nematode genomes have been sequenced is the possibility of identifying shared regulatory elements by phylogenetic foot printing (Stein *et al,* 2003). In order to identify any potential shared regulatory elements for the *daf-2* genes the DNA sequences of the intergenic space 5'of all four *daf-2* genes was aligned using the LAGAN alignment algorithm on the mVISTA server (http://genome.lbl.gov/vista/index.shtml) (Frazer *et al.*, 2004). The 5' and 3' flanking genes for *C. elegans* and *C. briggsae* have already been identified (Table 4.3.6).

In order to the find the 5' flanking genes for *C. remanei* and *B. malayi* the entire upstream sequence available from the contigs on which their *daf-2* genes are located was subjected to BLASTX (Gish and States, 1993) search of the *C. elegans* protein database (Wormpep, http://www.wormbase.org/db/searches/blat). In the case of *C. remanei* this represented 9695 bp and for *B. malayi* it was 2666 bp. Neither of these regions returned any hits to the Wormpep database, suggesting that the 5' flanking genes were further upstream than the sequence available. For *B. malayi* it was not possible to identify any sequence further upstream than the ~2.7 kb available on the 14419 contig, as the contigs are still being assembled into a larger scaffold.

There is, however, a *C. remanei* supercontig scaffold. A ~2.5Mb supercontig named Contig5 was identified by blasting the predicted *Cr-daf-2* genomic sequence across the supercontig assembly of the *C. remanei* genome. A 30 kb fragment of DNA sequence from this contig that was upstream of the predicted start codon of the *Cr-daf-2* gene was then blasted against the Wormpep database; several genes were identified as being on this fragment. The closest to the start codon of *Cr-daf-2* was the *C. remanei* orthologue of Y55D5A.6 (*Cr*-Y55D5A.6), a cuticulin precursor. The 5' intergenic space between the stop codon of the *Cr*-Y55D5A.6 gene and the start codon of the *Cr-daf-2* gene was 15032 bp.

The alignment of the 5' intergenic space from the four different nematodes, using a sequence identity threshold of 70% and a sliding window of 20 bp, failed to find any conserved regions between all four species. Similarly an mVISTA alignment of just the 5' intergenic space from the three *Caenorhabditis* species using the same threshold as before yielded small segments of sequence conservation between any pair of species,

mostly runs of a single nucleotide, but nothing was conserved between all three (data not shown).

In order to ensure that there were no regulatory elements within the intronic sequence of the *daf-2* genes or potentially the 3' intergenic space, an alignment of the entire genomic sequence including the 5' and 3' intergenic spaces using the S-LAGAN algorithm was performed. This programme takes in to account the possibility of rearrangements within the sequences being aligned. The 3' flanking genes were identified using the same procedure as the 5' flanking genes. This search only returned conserved regions between all four genes that represented exonic sequence; there were no conserved sequences within the intronic sequence or 3' intergenic space of the four genes. The inability to identify any conserved elements raises two possibilities. Firstly, that there has been rapid evolution of the regulatory elements of the *daf-2* genes in these different nematodes, resulting in highly divergent motifs. The other possibility is that motifs may be very short, and that the search criteria were not sensitive enough to identify them.

Table 4.3.6 Genes flanking *daf-2* in four nematode species

5' flanking gene	5' intergenic space	*daf-2*	3' intergenic space	3' flanking gene
Y55D5A.6	14023 bp	*daf-2*	5863 bp	Y55D5A.4
CBG1733 (orthologue of R10F2.1)	10155 bp	*Cb-daf-2*	5710 bp	CBG15731 (No *C. elegans* orthologue)
Cr-Y55D5A.6	15032 bp	*Cr-daf-2*	12390 bp	*Cr*-T20B6.3
Not known	>2666 bp	*Bm-daf-2*	6961 bp	*Bm-ceh-20*

4.3.7 Comparison of the amino acid sequence of DAF-2 proteins

The predicted amino acid sequences of the four DAF-2 proteins were compared. Of particular interest was the degree of conservation of the N- and C-terminal extensions. The lengths of the DAF-2 proteins from the four different species are as follows: DAF-2

is 1846aa long, *Cb*-DAF-2 is 1807aa, *Cr*-DAF-2 is 1816aa and *Bm*-DAF-2 is 1456aa. The overall percentage of identity and similarity between all four proteins and *h*IR is outlined in Table 4.3.7. The main reason why *Bm*-DAF-2 is so much shorter than the three *Caenorhabditis* DAF-2 proteins is that it lacks an N-terminal extension (Table 4.3.8 and Figure 4.3). *Bm*-DAF-2 also has a much shorter C-terminal extension than the other nematode DAF-2 receptors. However, the N-terminal extensions of *Cb*-DAF-2 and *Cr*-DAF-2 are both slightly shorter than that of DAF-2.

Surprisingly the percentage identity and similarity between *Cb*-DAF-2 and *Cr*-DAF-2 is slightly lower than that of DAF-2 and *Cr*-DAF-2 (Table 4.3.8), even though *C. briggsae* and *C. remanei* are likely sister species and share a more recent ancestor than with *C. elegans*. This may be due to the changes in the coding sequence caused by the gain of ten additional introns in the *Cb-daf-2* gene. It should be noted that the percentage sequence identity and similarity between DAF-2 and *h*IR given below is lower than that stated in section 4.1.1. The values given in that section are those calculated by Kimura *et al* (1997) and do not include the N- and C-terminal extensions of DAF-2. Those given in Table 4.3.8 do include the extensions so that the level of overall conservation between *Caenorhabditis* DAF-2 proteins can be gauged.

Table 4.3.7 The % identity and similarity of the DAF-2 proteins

	DAF-2 (Identity: Similarity)	*Cb*-DAF-2 (Identity: Similarity)	*Cr*-DAF-2 (Identity: Similarity)	*Bm*-DAF-2 (Identity: Similarity)	*h*IR* (Identity: Similarity)
DAF-2		50.2%: 65.5%	55%: 70.5%	30.2%: 43.6%	24.5%: 36.7%
***Cb*-DAF-2**	50.2%: 65.5%		54.9%: 67.4%	31.0%: 45.2%	26.2%: 39.1%
***Cr*-DAF-2**	55.0%: 70.5%	54.9%: 67.4%		30.5%: 44.2%	25.7%: 37.7%
***Bm*-DAF-2**	30.2%: 43.6%	31.0%: 45.2%	30.5%: 44.2%		31.5%: 46.2%
h*IR	24.5%: 36.7%	26.2%: 39.1%	25.7%: 37.7%	31.5%: 46.2%	

The percentage of shared identity and similarity (Identity: Similarity) between the four different DAF-2 proteins and *h*IR across the whole of their length.

* The % of sequence identity and similarity are smaller than those previously reported by Kimura *et al* (1997), as these alignments include the N- and C-terminal extensions.

Table 4.3.8 The position of domains within each DAF-2 protein

Protein / Domain/ Feature	DAF-2	*Cb*-DAF-2	*Cr*-DAF-2	*Bm*-DAF-2
N-terminal Extension	1-140	1-122	1-109	-
Signal peptide	122-142	103-123	92-110	1-16
L1	144-329	125-318	112-307	20-181
Cysteine Rich	330-490	319-474	308-463	182-336
L2	491-650	475-637	464-623	337-499
Fn0	651-775	638-758	624-748	500-619
Fn1	776-1072	759-1032	749-1028	620-861
Fn2	1073-1163	1033-1120	1029-1119	862-954
Transmembrane	1185-1207	1142-1164	1138-1160	980-1002
Tyrosine Kinase	1230-1540	1187-1498	1186-1495	1058-1342
Putative SH2-binding site	-	1512-1515*	1510-1513*	1347-1350
Putative SH2-binding site	1626-1629	1592-1595	1588-1591	1397-1400
Putative SH2-binding site	1649-1652	1616-1619	1612-1615	-
Putative SH2-binding site	1713-1716	1679-1682	1673-1676	-
Putative SH2-binding site	1800-1803	1758-1761	1768-1771	-

Position of domains and motifs in each of the different DAF-2 protein amino acid sequences from their N- to C- terminus.

* Additional phosphotyrosines that may be a putative SH2-binding site predicted by NetPhos server, no tyrosine residue at equivalent position in DAF-2 sequence.

4.3.8 Conservation of N- and C-terminal extensions in the *Caenorhabditis* genus

Like DAF-2, the *Cb*-DAF-2 and *Cr*-DAF-2 proteins also have an N-terminal extension in front of their predicted signal peptides, however, it is shorter in *Cb*-DAF-2 and shorter still in *Cr*-DAF-2 (Table 4.3.3). These extensions seem to be unique to *Caenorhabditis*; even though dInR has a much longer N-terminal extension it shares very little homology with the *Caenorhabditis* sequences apart from the predicted signal peptide (Table 4.3.9 and Figure 4.3). The *Caenorhabditis* DAF-2 N-termini share ~30% identity and ~46% similarity. A pair of putative endoprotease cleavage sites, N-terminal of the putative signal peptide, is conserved in all three DAF-2 N-termini. This suggests that this region is cleaved off at some point during the processing of the proreceptor of each of these proteins. However, it is difficult to speculate on the possible function of these cleaved peptides, as they do not have homology to any known proteins. Interestingly, the putative endoprotease cleavage sites are located at the boundary of exon 3 in both *daf-2* and *Cb-daf-2*, and exon 2 of *Cr-daf-2*. The proceeding exon in each gene (exon 4 in *daf-2* and *Cb-daf-2*, and exon 3 of *Cr-daf-2*) encodes the predicted signal peptide (Figure 4.4). The intervening intron in all three genes is in phase 0, which suggests that N-terminal extension was gained by 'exon-shuffling' (Gilbert, 1978).

The three *Caenorhabditis* DAF-2 proteins also have very similar C-terminal extensions, with the DAF-2 and *Cr*-DAF-2 extensions sharing slightly higher conservation of sequence identity and similarity (Table 4.3.10). All four putative SH2-binding sites identified by Kimura *et al* (1997) are present in both the *Cb*-DAF-2 and *Cr*-DAF-2 C-terminal extensions (Figure 4.5). Interestingly a screen for potential phosphotyrosines in the C-terminal extensions of all three *Caenorhabditis* DAF-2 proteins using the NetPhos (Blom *et al.*, 1999) server (http://www.cbs.dtu.dk/services/NetPhos/) revealed an additional putative SH2-binding site conserved in both *Cb*-DAF-2 and *Cr*-DAF-2 but not present in DAF-2 (Table 4.3.8 and Figure 4.5). It is unclear from the alignment if this tyrosine residue represents a gain in the common ancestor of *C. briggsae* and *C. remanei* or a loss in the DAF-2 sequence.

An interesting feature of the C-terminal extension in all three *Caenorhabditis* DAF-2 proteins is that it is composed of three repeats, each of which is encoded by an individual exon in the 3' end of the gene (Figure 4.6). The position of introns between these exons is also conserved between these three *daf-2* genes. This suggests that

elongation of the C-terminus in the *Caenorhabditis* DAF-2 proteins occurred via exon shuffling of an original final exon in an ancestor of these three species. This would also explain why the *Bm*-DAF-2 C-terminus is shorter than that of the *Caenorhabditis* DAF-2 proteins. Thus, exon shuffling has been an important process in the evolution of the *daf-2* gene with the *Caenorhabditis* genus.

Table 4.3.9 The % identity and similarity of the N- terminal extensions

	DAF-2 (Identity: Similarity)	***Cb*-DAF-2 (Identity: Similarity)**	***Cr*-DAF-2 (Identity: Similarity)**	**dInR (Identity: Similarity)**
DAF-2		32.9%: 46.2%	34.2%: 46.6%	10.9%: 18.6%
***Cb*-DAF-2**	32.9%: 46.2%		32.9%: 40.1%	6.7%: 10.7%
***Cr*-DAF-2**	34.2%: 46.6%	32.9%: 40.1%		7.4%: 10.2%
dInR	10.9%: 18.6%	6.7%: 10.7%	7.4%: 10.2%	

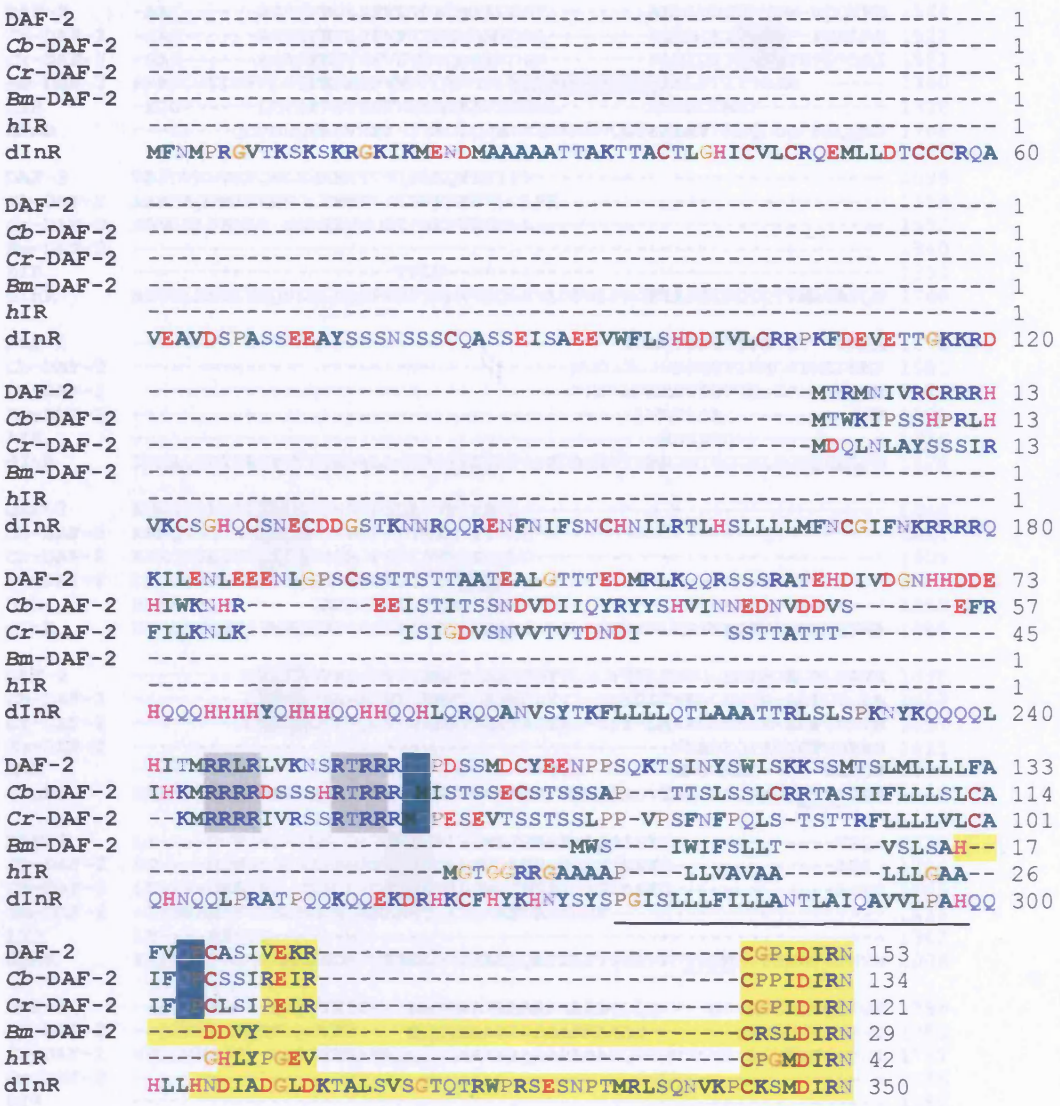
The percentage of shared identity and similarity (Identity: Similarity) between the *Caenorhabditis* DAF-2 proteins and dInR across their N-terminal extensions. The *Bm*-DAF-2 protein was excluded from this analysis, as it has no N-terminal extension

Table 4.3.10 The % identity and similarity of the C- terminal extensions

	DAF-2 (Identity: Similarity)	***Cb*-DAF-2 (Identity: Similarity)**	***Cr*-DAF-2 (Identity: Similarity)**	***Bm*-DAF-2 (Identity: Similarity)**	**dInR (Identity: Similarity)**
DAF-2		53.8%: 67.9%	60.1%: 72.5%	13.3%: 18.6%	12.3%: 21.3%
***Cb*-DAF-2**	53.8%: 67.9%		59.5%: 67.9%	13.3%: 19.6%	15.3%: 26.5%
***Cr*-DAF-2**	60.1%: 72.5%	59.5%: 67.9%		10.5%: 15.1%	15.7%: 24.6%
***Bm*-DAF-2**	13.3%: 18.6%	13.3%: 19.6%	10.5%: 15.1%		6.4%: 11.3%
dInR	12.3%: 21.3%	15.3%: 26.5%	15.7%: 24.6%	6.4%: 11.3%	

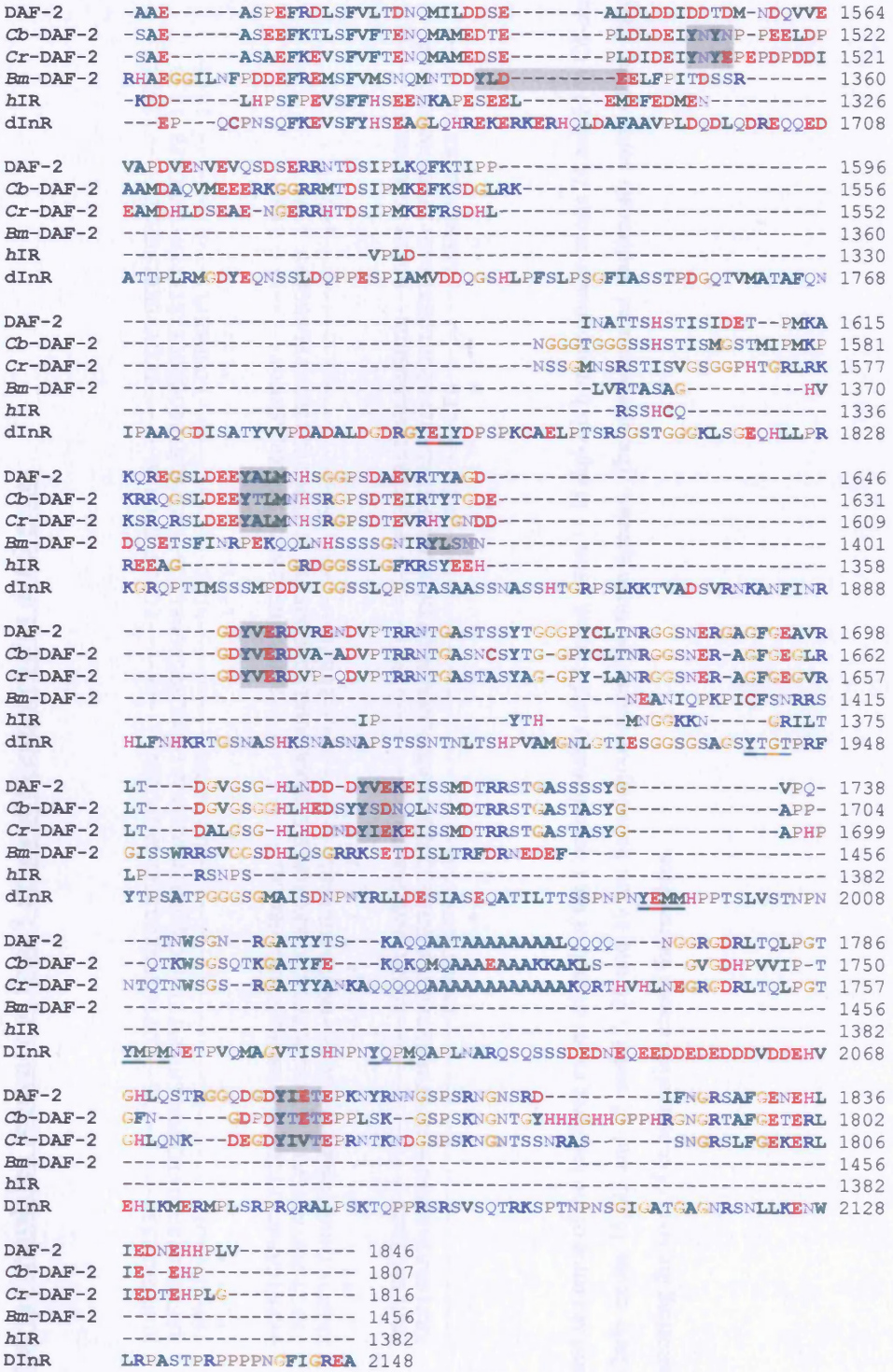
The percentage of shared identity and similarity (Identity: Similarity) between the four different DAF-2 proteins and dInR across their C-terminal extensions..

Figure 4.4 Alignment of N-terminus of DAF-2 proteins



Alignment of N-termini of the four different DAF-2 proteins and *h*IR and dInR. The sequences shaded in grey are putative endoprotease cleavage sites. The region underlined represents the putative signal peptide of each protein, as predicted by the SignalP program (Dyrlov Bendtsen *et al.*, 2004). The sequences highlighted in yellow are part of the L1 domain, the residues immediately N-terminal of this point are the predicted cleavage sites for the signal peptide, and were used as anchoring residues in the alignment of the N-termini. The sequences highlighted in blue are residues interrupted by phase 0 introns.

Figure 4.5 Alignment of C-terminus of DAF-2 proteins



Alignment of C-termini of the four different DAF-2 proteins and *h*IR and dInR. Residues shaded in grey are the putative SH2-binding sites predicted by the NetPhos server (Blom *et al.*, 1999) for each of the DAF-2 proteins. The residues that are underlined in the dInR C-terminal extension are predicted NetPhos predicted SH2-binding sites.

Figure 4.6 Alignment of the translated sequence of final 3 exons of *Caenorhabditis daf-2* genes

```
DAF-2 exon 15        RRNTGASTSSY-------TG--GGPYCLTNRGGSNERGAGFGEAV----------RLTDGVGSGHLN------DDDYVEKEISSMDT----- 1723
DAF-2 exon 16        RRSTGASSSSYGVPQTNWSGNRGATYYTSKAQQAATAAAAAAAALQQQQNGGRGDRLTQLPGTGHLQSTRGGQDGDYIETEPKNYRNNGSPS 1815
DAF-2 exon 17        -RN-GNSRDIF-------NG-----------------RSAFGENE----------HLIE--------------DNEHHPLV----------- 1846
                      *. * * . :       .*                  :. .            :* :              *.::

Cb-DAF-2 exon 24     -RRNTGASNCSYTGGPYCLTNRGGSNERAGFGEGLRL-----------TDGVGSGGHLHEDSYYSDNQLNSMDT-------- 1689
Cb-DAF-2 exon 25     -RRSTGASTASYGAPPQTKWSGSQTKGATYFEKQKQMQAAAEAAAKKAKLSGVGDHPVVIPTGFNGDPDYTETEPPLSKGSPS 1771
Cb-DAF-2 exon 26     KNGNTGYHHHGHHGPP----HRG--NGRTAFGETERL-----------IEEHP---------------------------- 1807
                      . .**    .: . *      . ::  : * :  ::            .

Cr-DAF-2 exon 14     RRNTGASTASY-------------AGGPYLANRGGSNERAG-----------------FGEGVRLTDALGSGHLH--DDNDYIEKEISSMDT----- 1683
Cr-DAF-2 exon 15     RRSTGASTASYGAPHPNTQTNWSGSRGATYYANKAQQQQQAAAAAAAAAAAAKQRTHVHLNEGRGDRLTQLPGTGHLQNKDEGDYIVTEPRNTKNDGSPS 1783
Cr-DAF-2 exon 16     ------------------------KNGNTSSNRASSN----------------------GRS-----LFG--------EKERLIEDTEHPLG----- 1816
                                          .. :*:. .:                        *..       *       ::: :
```

An alignment of the translated sequence of the last three exons of each of the *Caenorhabditis daf-2* genes, exon 15 in *daf-2* is homologous to exons 24 and 14 in *Cb-daf-2* and *Cr-daf-2* respectively. Similarly exons 16, 25 and 15 and 17, 26 and 16 are homologous across the three species. The symbols used underneath some of the residue are standard ClustalW codes indicating the level of conservation across the residues.

4.4 Discussion

In this chapter I determined the sequence of two new nematode insulin/IGF-1 receptors and created a model for a third. Comparing the four nematode DAF-2 receptors has revealed a number of features of nematode *daf-2* gene and protein structure, and given insight into the evolution of this gene within the nematode phylum.

4.4.1 *daf-2* gene structure and intron losses and gains

Intron loss has played a major part in the evolution of the nematodes especially the *Caenorhabditis* genus (Cho *et al.*, 2004; Kiontke *et al.*, 2004). The *daf-2* genes from four different species of nematode *B. malayi*, *C. elegans*, *C. briggsae* and *C. remanei* generally fit this pattern of intron loss. By comparing introns that are conserved between these four species and several outgroup species it becomes apparent that the *Caenorhabditis* species have lost at least five of the introns present in the *Bm-daf-2* gene. By contrast, only one intron present in the *Caenorhabditis* has been lost from *Bm-daf-2*. More surprisingly the *daf-2* genes in the *Caenorhabditis* species have at least five intron gains at the 5' and 3' ends of the gene compared to *Bm-daf-2*. I postulate that these additional introns explain the origin of the N-terminal extension and the increased length of the C-terminal extension of the *Caenorhabditis* DAF-2 proteins, most likely via a process of exon shuffling (Gilbert, 1978).

In addition to the intron gains shared with *C. elegans* and *C. remanei*, the *Cb-daf-2* gene has gained an additional 10 introns. Examples of intron gains are rare, especially within the *Caenorhabditis* genus (Coghlan and Wolfe, 2004), and the mechanism by which they occur is not yet known. One possible mechanism of intron gain is the reverse splicing of pre-existing introns back into an alternative site in the mRNA, which may then recombine with the genomic copy of the gene to introduce an additional intron. There are several possible examples of this in the *C. elegans* and *C. briggsae* genome (Coghlan and Wolfe, 2004). In these cases the gained introns had very high sequence similarity to other introns either from within the gene or other genes in the genome.

My own analysis of the gained introns in *Cb-daf-2* does not support the idea that they are related to other introns in the *Cb-daf-2* gene as none have a high degree of sequence similarity to existing introns (data not shown). However this analysis may be complicated by the fact that the intronic sequence of the gene is very repetitive (when analysed with the *C. elegans* repeat library). A search for sequence similarity amongst other introns across the *C. briggsae* genome may yield the origin of the gained introns.

Currently the genome of the unnamed *Caenorhabditis* species CB5161 (strain PB2801) is being prepared for sequencing. The position of this species with regards to *C. briggsae* and *C. remanei*, to both of whom it is highly related cannot currently be resolved using available molecular phylogenies (Cho *et al.*, 2004; Kiontke *et al.*, 2004). It would be both interesting and beneficial to either model or sequence the *daf-2* gene from this species. If the *daf-2* gene from this species has gene structure more similar to *Cb-daf-2* than *Cr-daf-2* it would suggest a more ancient origin for the additional introns found in *Cb-daf-2*. This could explain the lack of sequence similarity between the gained introns and conserved introns within the *Cb-daf-2* gene, as well as the accumulation of repetitive sequence in these introns. If however the *daf-2* gene structure from CB5161 is more similar to *Cr-daf-2*, then this would raise the more difficult question of how the *Cb-daf-2* gene had gained 10 introns, given that gains are rare.

4.4.2 Lack of a conserved *daf-2* promoter

Analysis of the intergenic sequence 5' of each of the four different *daf-2* genes did not identify any conserved elements that might represent a *daf-2* promoter. This may be due to two reasons; firstly there is a large evolutionary distance between these four species, the estimated date of divergence between *B. malayi* and *C. elegans* is 450 MYA (M. L. Blaxter, personal communication), which may be too large to reliably identify conserved regulatory sequences. One should note that only ~2.6 kb of upstream sequence was available for the *Bm-daf-2* gene; potentially a more conserved element could lie further upstream.

The estimated date of divergence between *C. elegans* and *C. briggsae* is 80-110 MYA (Stein *et al*, 2003) and a similar date can be estimated for the divergence of *C. remanei* and *C. elegans*, given that *C. remanei* and *C. briggsae* are sister species and share a common ancestor. Even at this evolutionary distance conserved regulatory

elements between *C. elegans* and *C. briggsae* have been found for several different genes (GuhaThakurta *et al.*, 2004; Petalcorin *et al.*, 2005). One possibility is that *C. remanei* and *C. briggsae* may share a conserved regulatory element that has diverged from that of *C. elegans*. This possibility may soon be testable as two additional *Caenorhabditis* genomes are expected to become available by the end of this year (2005). These are the genomes of unnamed species CB5161 and *C. japonica*, which is thought to be the outgroup species to the Elegans group of worms (*C. elegans, C. remanei, C. briggsae* and CB5161).

A second possible reason for the lack of a conserved regulatory element may be due to the fact that the *daf-2* gene seems to lie in a region that is a hotspot for intrachromosomal rearrangement. The flanking genes on either side of the *daf-2* gene in each species are different in most cases (Table 4.3.6). Take the 5' flanking genes: *daf-2* and *Cr-daf-2* are both flanked by a cuticulin precursor gene (Y55D5A.6 and *Cr*-Y55D5A.6). But the *Cb-daf-2* gene is flanked by CBG15733, an orthologue of the *C. elegans* R10F2.1 gene. The 5' flanking gene of *Bm-daf-2* is unknown as only a small amount of 5' intergenic sequence is available.

In the case of the 3' flanking genes, none of these are conserved in the same location. The *daf-2* 3' flanking gene is Y55D5A.4, while the 3' flanking gene for *Cb-daf-2* is CBG15731, a novel gene with regards to *C. elegans*. A TBLASTN search of the *C. remanei* genome with CBG15731 does identify a probable orthologue, but is located on a different contig to that of *Cr-daf-2*. The 3' flanking gene of *Cr-daf-2* is *Cr*-T20B6.3, the orthologue of the *C. elegans* T20B6.3 gene and the 3' flanking gene of *Bm-daf-2* is *Bm-ceh-20*, the orthologue of *ceh-20*.

All of the *C. elegans* orthologues of the different flanking genes, except CBG15731, which has no orthologue, are located on the same linkage group (III) as *daf-2*. This suggests that although the microsynteny of genes immediately around *daf-2* is not conserved between the four different species there is some macroscopic conservation of gene order. Possibly, there are *daf-2* regulatory elements (conserved or otherwise) further upstream, away from this rearrangement prone hotspot.

4.4.3 Conservation of the N-terminal and C-terminal extensions in DAF-2 proteins

A major aim of this comparative analysis of different nematode DAF-2 receptors was to gain further insight into the extensions found at both termini of the *C. elegans* DAF-2 protein. The N-terminal extension of the DAF-2 protein, predicted by Kimura *et al* (1997) when the *daf-2* gene was originally sequenced, is also present in the predicted *Cb*-DAF-2 and *Cr*-DAF-2 proteins, but not in the *Bm*-DAF-2 protein (Table 4.3.8 and Figure 4.3 and 4.4). Thus, this extension is not universal to nematodes, but is well conserved in the *Caenorhabditis* genus (Table 4.3.9). This implies that this extension has functional importance within this genus. The N-terminal extension in all three *Caenorhabditis* DAF-2 proteins may be endoproteolytically cleaved (Figure 4.4). However, it is difficult to postulate a function for this cleaved peptide, as it has no homology to any known protein domain.

The presence of the N-terminal extension outside the *Caenorhabditis* genus currently remains unknown. It will be possible to model or sequence the *Cj-daf-2* gene when the *C. japonica* genome sequence becomes available later this year. If the N-terminal extension were a unique feature of the Elegans group within the *Caenorhabditis* genus, then the *Cj*-DAF-2 protein would not be predicted to have the N-terminal extension.

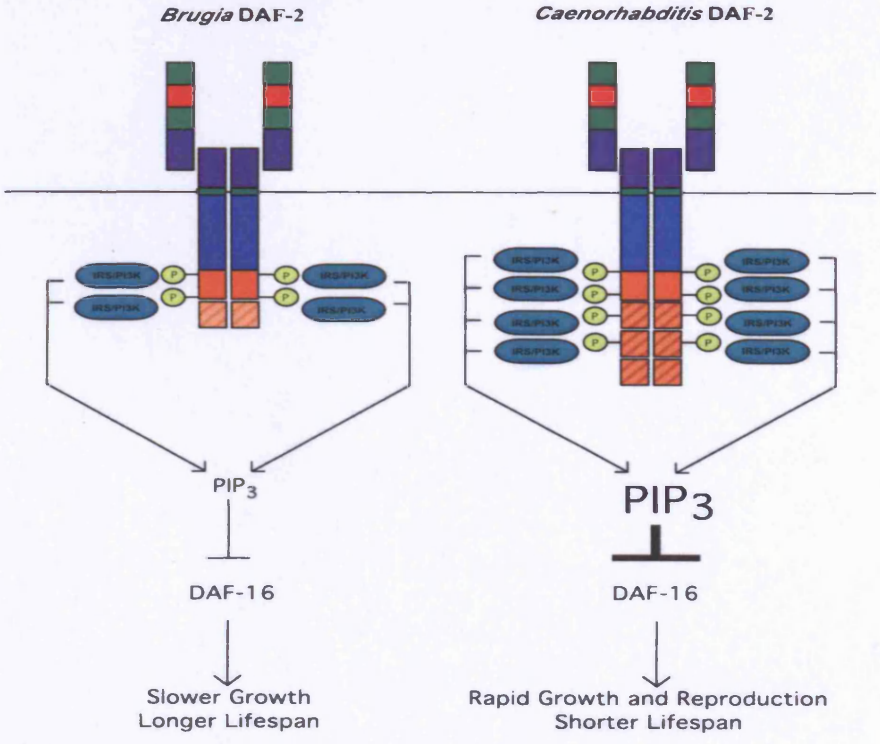
A more interesting possibility is that the extension is a feature of the *Heterorhabditis* group of nematodes, which includes the *Caenorhabditis* and *Strongyloides* genii, but not *Brugia* (Blaxter *et al.*, 1998). The sequence of a fifth *daf-2* gene (*irlk-1*) from the human parasitic nematode *Strongyloides stercoralis* should be available shortly (H. C. Massey and J. B. Lok, personal communication). If the IRLK-1 protein has a N-terminal extension like the *Caenorhabditis* DAF-2 proteins, then such a structure must have been present in the last common ancestor of these worms approximately 320 MYA (M. L. Blaxter, personal communication).

All four DAF-2 proteins have C-terminal extensions that contain putative phosphotyrosine residues that may act as SH2-binding sites. The C-terminal extension of the three *Caenorhabditis* DAF-2 proteins is longer than that of *Bm*-DAF-2. The *Bm*-DAF-2 has two putative SH2-binding sites, while DAF-2 has four (Kimura *et al*, 1997) and *Cb*-DAF-2 and *Cr*-DAF-2 have five (Table 4.3.8). At present it is not possible to tell whether this due to the loss of a putative site in DAF-2 or gain in the other two proteins. Similar C-terminal extensions are found in the majority of invertebrate IRRs, with the sponges being notable exceptions (Skorokhod *et al.*, 1999). These extensions

are thought to directly interact with substrates, such as the PI3 kinase, unlike the vertebrate IRRs, which rely on adaptor proteins such as the IRS family.

Many aspects of *C. elegans* biology imply that this species has evolved for rapid population growth. The extension of the C-terminus of DAF-2 by exon shuffling may be such an adaptation. The number of substrates that can bind to the C-terminus of an IRR may affect the level of signal generation by that receptor. This being so, then the amplification of the C-terminus in DAF-2, and the other *Caenorhabditis* DAF-2 proteins, may in turn lead to an amplification of signal generation by the receptor. Thus, one might predict that the binding of a single insulin-like ligand to a DAF-2 receptor molecule would result in more PIP3 generation than a single *Bm*-DAF-2 receptor, whose C-terminal extension only has two predicted SH2-binding sites. Increased levels of PIP3 would lead to greater negative regulation of DAF-16, which would promote entry into reproductive development.

Figure 4.7 C-terminal extension as a means for promoting rapid growth



Hypothetical model of how the extended C-terminal extension (orange segments) of the *Caenorhabditis* DAF-2 proteins contributes to their rapid lifecycle. Binding of a single insulin-like ligand to the *Caenorhabditis* DAF-2 receptors allows for the activation of a greater number of effector proteins than the binding of a single molecule to the *Bm*-DAF-2 receptor. This in turn leads to the generation of a greater amount of PIP3 leading to greater suppression of DAF-16, hence more rapid development and shorter lifespan.

Given the conservation of the N- and C-terminal extensions within the *Caenorhabditis* genus, it is interesting that no mutations (out of 23 sequenced alleles) have been identified in the DAF-2 protein that affect either of these regions (see Chapter 5). However, there is some evidence for a functional role for the C-terminal extension in DAF-2 (see Chapter 6). It may be possible to test for a functional role for the N-terminal extension with mutations that affect the cleavage site, which might result in a phenotype if their retention interferes with wild-type DAF-2 activity. A similar effect is seen in the semi-dominant *daf-28(sa191)* allele, which has a mutation that affects the endoproteolytic cleavage site of the DAF-28 protein. The mutant protein behaves as a poisonous ligand of the DAF-2 receptor and causes a Daf-c phenotype (Li *et al.*, 2003). Overexpression of a *daf-28(sa191)* cDNA construct in a wild-type background phenocopies the effect of the homozygous mutant. It might be informative to remove the cleavage sites in the N-terminal extension of DAF-2, through i*n vitro* mutagenesis, and then overexpress this construct in wild-type worms.

Chapter 5

Sequence analysis of *daf-2* mutant alleles

5.1 Introduction

In this chapter I aim to test the correlation between the location of a mutation in DAF-2 and the phenotypic class of the allele bearing that mutation. I also explore the use of homology modelling of the DAF-2 receptor at secondary, tertiary and quaternary levels of protein structure and its ability to generate hypotheses as to the molecular nature of the difference that leads to two classes of *daf-2* allele.

5.1.1 Possible correlation between location of mutation and phenotypic class

Previous phenotypic analysis of 15 temperature-sensitive (ts) alleles of the *daf-2* gene revealed the existence of two overlapping pleiotropic classes (Gems *et al.*, 1998). Several of the *daf-2* alleles used in that study had had their mutations identified by Kimura *et al* (1997). These were *e1365* (A580T) and *e1368* (S573L), both of which are class 1, *e1370* (P1465S) and *e1391* (P1434L), which are class 2. A further four alleles of *daf-2* were also sequenced in Kimura *et al* (1997), one of these, *mg43* (C410Y, P470L), is a non-conditional allele. The three other alleles were partially characterised in Gems *et al* (1998) and assigned to a class: *sa229* (D648N) is class 1 and *sa187* (C469S) and *sa219* (D1374N) both appear to be class 2.

Scott *et al* (2002) have also sequenced several alleles of *daf-2* that had been characterised in Gems *et al* (1998). Two of these were *m579* (R437C) and *m596* (G547S) both of which were reported as being class 2. Scott *et al* (2002) had also reported that the *e979* allele, which is class 2, had a G383E mutation. However, the *m41* allele of *daf-2*, which is class 1, had previously been reported to also have a G383E substitution (Yu and Larsen, 2001) and this has been confirmed in our lab (D. McCulloch and D. Gems, unpublished results).

To test for a possible correlation between the location of a mutation and the phenotypic class of the allele bearing it I sequenced additional alleles that had been assigned a class by Gems *et al* (1998). I also resequenced both the *m41* and *e979* alleles, in order to clarify the nature of the mutation in each allele. I also sequenced several non-conditional alleles of *daf-2*, including *m65* a candidate null allele, to gain further insight into DAF-2 function.

5.1.2 The value of homology modelling of the DAF-2 receptor

Over 50 naturally occurring mutations in *h*IR have been identified and characterised biologically (see Taylor *et al.*, 1992 and Krook *et al.*, 1997a for detailed reviews) and four have also been found in the *h*IGF-1R (Abuzzahab *et al.*, 2003; Kawashima *et al.*, 2005). In addition, many mutagenesis studies have been performed on both receptors to identify residues involved in ligand binding (Mynarcik *et al.*, 1997) and those associated with signal transduction (Strack *et al.*, 1997; Yam *et al.*, 2001).

Crystal structures are available for the first three domains (L1, CR and L2) of the α-subunit of the *h*IGF-1R (Garrett *et al.*, 1998) and the TK domains of both the *h*IR (Hubbard *et al.*, 1994) and *h*IGF-1R (Favelyukis *et al.*, 2001). In addition, a low-resolution quaternary structure of *h*IR bound to insulin has been obtained through scanning electron microscopy (Luo *et al.*, 1999). This quaternary structure has been combined with the available crystal structures from above to generate an atomic-scale model of the quaternary structure of *h*IR that generates a possible mechanism for transmembrane signal transduction through the receptor (Ottensmeyer *et al.*, 2000).

In order to gain insight into the mode of action of different DAF-2 mutations I performed a comparative analysis of both new and previously sequenced alleles of *daf-2* with mutations in *h*IR and *h*IGF-1R. I also used homology modelling of the DAF-2 receptor to gain insight into the effects of mutations on the secondary and tertiary structure of the receptor. This work was carried out in close collaboration with Acely Garza-Garcia, a PhD student in the Department of Biochemistry at UCL and her supervisor Dr. Paul Driscoll. We then attempted to gauge the impact of the structural changes caused by our mutations on signal transduction through the DAF-2 receptor by collaborating with Prof. Peter Ottensmeyer at the University of Toronto, who placed our mutations into a quaternary model of DAF-2.

5.2 Materials and methods

5.2.1 Strains

The following strains carried different alleles of *daf-2* that had their cDNA sequenced in order to determine the nature of the mutation they carried; CF1844 *fer-15(b26) II; daf-2(mu150) III; fem-1(hc17) IV*, DR608 *daf-2(m212) III*, DR1567 *daf-2(m577) III*, DR1568 *daf-2(e1371) III*, DR1573 *daf-2(e1369) III*, DR1942 *daf-2(e979) III*, GA156 *daf-2(m631) III; daf-18(nr2037) IV*, GA158 *daf-16(mgDF50) I, daf-2(m65) III*, GA310 *daf-16(mgDf50)*; *daf-2(tm1236) III*, JT193 *daf-2(sa193) III* and JT6723 *daf-2(sa223) III/ qC1*[*dpy-19(e1259), glp-1(q339)*] *III*. Cynthia Kenyon provided the CF1844 strain and James H. Thomas provided the JT193 and JT6723 strains.

5.2.2 *daf-2* cDNA primer design

Initial primers for the amplification of cDNA fragments from the different *daf-2* alleles were designed based on the *daf-2* cDNA sequence of Kimura *et al* (1997). Primers were designed using the Primer3 WWW interface (Rozen and Skaletsky, 2000) (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Primers were generally designed to amplify overlapping fragments 350bp-500bp in size.

5.2.3 Sequencing of *daf-2* alleles

Total RNA from mixed stage worms was isolated using the procedure described in chapter 3.2.1 for each of the *daf-2* mutants used, except *daf-2(sa223)*. This allele is non-conditional and was not available as a double mutant with a suppressor of its Daf-c phenotype. So to generate RNA for *daf-2(sa223)* homozygotes I grew the JT6723 strain on sixteen 60mm plates seeded with OP50 at 20°C. I then washed the worms off the plates and into a 15ml centrifuge tube using 10ml of M9. The worms were then centrifuged at 1000rpm for 1min. the supernatant was removed and the worms were resuspended in 10ml of 1% SDS solution (W/V). The animals were incubated in the

SDS solution for 1 hour, to kill all animals except dauers and unhatched eggs, before they were centrifuged at 1000rpm for 1min. The supernatant was removed and the pellet was resuspended in 10ml of M9. The worms were then centrifuged again at 1000rpm for 1min. The pellet was washed a further two times with M9, using the procedure described above. After the final wash the worms were resuspended in 200µl of M9 and transferred on to a fresh seeded plate and left overnight at 20°C.

On the following day the worms were washed of the plate using 3ml of M9 and transferred into a 15ml centrifuge tube, to which 7ml of 1% SDS solution was added. The worms were again incubated in the SDS solution for 1 hour before being washed using the procedure above. The second SDS wash was necessary to kill larvae that had hatched from the eggs that been retained from the first SDS wash. The dauers recovered from the SDS washing had their RNA extracted as described in Chapter 3.2.1. Methods for first strand cDNA synthesis from the RNA and amplification of and sequencing of overlapping fragments by PCR can be found in Chapter 3 (sections 3.2.2, 3.2.4 and 3.2.6). The sequences of the primers used for amplification and sequencing are listed in Table 5.2.1. The sequencing reads were analysed using the method described in Chapter 3.2.7.

Table 5.2.1 Primer pairs used in sequencing *daf-2* alleles

5' primer	5'-sequence-3'	3' primer	5'-sequence-3'	Product size
L	GCACACATCTGATCGTCGGATTC	F1R	TGGCGGGTTTTCCTCATAGCA	408bp
F2F	TGAGACGGCTTCGACTTGTCA	F2R	CGATGCATTTCTTCTTGTGCTT	388bp
F3F	TGCACAGTGGTGGAAGGTTCG	F3R	CATCGTTGATGGAAGAAGTGA	394bp
F4F	TCGAAATGGTGGTGTACGGATA	F4R	CGTGGCATGCTGTCGCATC	395bp
F5F	ATCGGACCGGGATGTGATGC	F5R	GCACTGTTTTGTTCGAGAGC	226bp
F6F	TGTGATGCTCACCTGTACCTTC	F6R	TCCTTCAACTCGGACGCCATTC	336bp
F7F	GCCGAAAATGCGTTGGCAAG	Ex9F1R	TGAACAGTGACTTTGCCTCAA	279bp
F8F	GGCTACCTGTTGGTACGTCAAT	F8R	CTCATCGATTCGTGGGACTT	447bp
F9F	GCGGACTCGGTCTTCTTTAG	F9R	TAGCGTCCGAATCGACTTGC	394bp
F10F	TGGCGACGCAGATGGTGTTG	F10R	GCAGCGCACGTTTTCTGCTG	391bp
F11F	CGATTGTAGCCGATAAGCCAGT	F11R	TGCTCGACGGCTTCTTCTTCG	384bp
F12F	CGGCGATCGAATCATCTGCA	Ex12rev	ATCCTTGTTCCCGGCACCGTAG	289bp
Ex11fwd	AGCAGAAAACGTGCGCTGCAA	F13R	TCCTTGTTCCCGGCACCGTA	384bp
F14F	GGCTCCAACGATCGCATCAA	F14R	TTCCACCATTCACTTCAGTCG	380bp
F15F	CATCGTGCGGGAGCATTGAA	F15R	GACGAGAAGCATGCCGAGAATGA	399bp
F16F	CTGGACCGGAAGCCGAATCC	F16R	TCGACGCTGGATCATCTACA	387bp
F17F	CTGATGGGTGATCGTTTCG	F17R	CATCGAACAGGCATCATACGC	482bp
F18F	TCGGGATGAGACTGTCAAGA	F18R	CCAGTGCTTCTGAATCGTCA	455bp
F19F	CCAGCTCGTTCATCTTCTAGC	F19R	TCCGAACTTCCGCATCACTCG	383bp
F20F	GCGAGAAGGATCGCTGGATG	F20R	CGTGGCACCACGATTTCCACT	394bp
F21F	TGAACGAGGAGCCGGTTTCG	R	CAGAGATTTGAAAAATGGGGTAAAA	615bp

Primer pairs used in sequencing the *daf-2* cDNA from different mutants and their wild-type product size. The fragments are overlapping and are listed from 5' to 3' in the cDNA sequence.

5.3 Results

To gain further insight into the molecular nature of class 1 and 2 alleles of *daf-2* and how this might lead to their differing phenotypes and genetic interactions I have sequenced several temperature-sensitive alleles of *daf-2*. I have also sequenced several non-conditional alleles of *daf-2*. This sequence data was then combined with homology modelling of the DAF-2 receptor and a comparative analysis of mutations that affect the *h*IR and *h*IGF-1R, to understand how these mutations affected the different levels of protein structure and how this might affect signalling through the receptor.

5.3.1 Selection of alleles for sequencing

The following temperature sensitive alleles of *daf-2* were initially chosen for sequencing: *e979*, *e1369*, *e1371*, *m212*, *m577*, *mu150* and *sa193*. All of these alleles, except *e979*, are class 1. The *e979* allele is interesting as it has the most severe Daf-c phenotype of the class 2 alleles at temperatures below 25°C. At 25°C the progeny of *e979* arrest either as embryos or L1s (Gems *et al.*, 1998; Vowels and Thomas, 1992). The alleles *e1369* and *m212* are both class 1, and yet have the most Daf-c phenotype of all ts alleles at 15°C forming 55% and 28% dauers, respectively (Gems *et al.*, 1998). The *e1371*, *mu150* and *sa193* alleles are all class 1A, which form transient dauers at 25°C that generally recover within 24 hours. The *mu150* allele has been extensively used by the Kenyon lab, and was included in this study at the request of Cynthia Kenyon (Garigan *et al.*, 2002; Hsu *et al.*, 2003; Murphy *et al.*, 2003). The *m577* allele is class 1C and is the most frequently used allele of *daf-2* in the Gems lab.

I also sequenced the following non-conditional alleles of *daf-2*: *m65*, *m631*, *sa223* and *tm1236*. The *m65* allele is a candidate null allele, which has an interesting embryonic and L1 arrest phenotype when placed in a *daf-16(-)* background (Larsen *et al.*, 1995)(M. Nanji and D. Gems, unpublished results). Epistasis analysis of the *m65*, *m631* and *m646*, a non-conditional allele bearing a nonsense mutation (K. Kimura and G. Ruvkun, personal communication) with *daf-16* and *daf-18* null alleles has revealed that *m631* has an intermediate level of suppression between *m65* and *m646* (M. Nanji and D. Gems, unpublished results, see Chapter 6 for more details). The *sa223* allele is

an interesting allele, since Daf-c is fully maternally rescued at 15°C, and rescued *sa223* adults are long lived and have greatly reduced fertility. The properties of *sa223*, especially the weakness of the Daf-c phenotype of the maternally rescued homozygotes, suggest weak reduction of *daf-2*A function coupled with strong reduction of *daf-2*B function (Gems *et al.*, 1998). The *tm1236* allele is a deletion mutant of *daf-2* that was isolated by the Japanese *C. elegans* knockout consortium (http://shigen.lab.nig.ac.jp/c.elegans/index.jsp) at my request. This allele was sent as a heterozygote over wild type and so placed over the *qC1* balancer and then outcrossed a further two times (see Chapter 6 for details)

5.3.2 Sequencing of *daf-2* alleles

The entire coding sequence of the *daf-2* transcript was amplified as a series of overlapping fragments using the primers listed in Table 5.2.1. These fragments were then sequenced on both strands. The sequencing reads for each strand were compared to the wild-type sequence reported by Kimura *et al* (1997) and the sequence of the *daf-2* predicted gene in WormBase. Fragments that contained identical nucleotide substitutions compared to the wild-type sequence were taken to be the mutation in the allele. This region was then amplified from genomic DNA isolated from the mutant worms and sequenced to confirm the presence of the lesion.

The sequence of the mutant cDNA was translated *in silico* and then compared to the wild-type protein sequences to determine the amino acid change (see Chapter 3.2.7). The two wild-type protein sequences differ in length by three amino acids (1846aa and 1843aa); the residues MTR at the N-terminus are not present on the WormBase sequence, which uses the second in-frame ATG codon as the start of translation instead of the first. I chose to use the numbering system of Kimura *et al* (1997) for the position of the mutations identified in this project. The Kimura *et al* (1997) protein sequence also differs from the WormBase protein sequence at two positions. The Kimura *et al* (1997) sequence has an arginine residue at position 838 and a glutamine residue as position 1313. Histidine and lysine residues occupy the equivalent positions in the WormBase sequence. All of my sequencing reads confirm the presence of the WormBase residues at these positions and not those of Kimura *et al* (1997). Below are the positions of the mutations that have identified in this project. A more detailed

exploration of each mutation and its predicted effect on the structure and signalling of the DAF-2 receptor will be dealt with in subsequent sections of this chapter.

The *e979* allele had previously been reported to have a G383E substitution in the CR domain of the DAF-2 receptor by Scott *et al* (2002). However, the *m41* allele has also been reported to have the same mutation (Yu and Larsen, 2001). The *m41* allele is class 1 and *e979* is class 2 and the Daf-c phenotype of *e979* is more severe than that of *m41* (Gems *et al.*, 1998). The *m41* is the only allele of *daf-2* in which dauer formation can be maternally rescued (Gems *et al.*, 1998). I sequenced and phenotyped our lab stock of *daf-2(e979)* and resequenced and phenotyped the *e979* allele used by Scott *et al* (2002). I found that the mutation in our lab stock of *e979* was C146Y, which affects the L1 domain. This stock phenotyped as previously reported in Gems *et al* (1998). Sequencing of the *e979* allele used by Scott *et al* (2002), obtained from the CGC, confirmed the G383E substitution in this strain. Phenotypic analysis confirmed that this strain was actually *m41* and not *e979*. Based on this information we asked the CGC to replace their stock of *daf-2(e979)* with our strain.

The *sa193* allele was found to have a missense mutation that caused an A580T substitution affecting the L2 domain of the receptor. Kimura *et al* (1997) had reported that the *e1365* allele also had the same mutation. However, Gems *et al* (1998) has reported that *sa193* is a class 1A allele of *daf-2*, whose dauers recover at 25°C after 24 hours, whereas *e1365* is a class 1C allele, whose dauers do not recover at this temperature. I resequenced and phenotyped the CB1365 strain from the CGC and sequenced and phenotyped our lab strain of *e1365*. I confirmed that the CB1365 strain contained the A580T mutation and that its dauers recovered at 25°C. Therefore *sa193* and *e1365*, two independently isolated alleles, have the same mutation. Our lab stock of *e1365* had a C1045Y mutation, which is the same mutation as the *m577* allele, and affects the Fn1 domain. Therefore the allele characterised as *e1365* in Gems *et al* (1998) was in fact *m577*.

The *e1369* allele was found to have no mutation in the CDS of the *daf-2* gene. This was confirmed by sequencing of full-length cDNA from two independent RNA extractions. This suggests that this allele might have a mutation in a regulatory element of the *daf-2* gene (see 5.3.9 for more details). The *e1371* allele has a C803Y mutation and the *m212* allele has a C883Y substitution, both affect the Fn1 domain. The *mu150* allele has a G682D substitution in the Fn0 domain. It should be noted that the *mu150*

allele was sequenced by Sindhuja Kadambi, a former third year research project student in the Gems lab, under my supervision.

The *sa223* allele has a missense mutation causing a R1430Q substitution within the TK domain of the receptor. This is the only allele, of those sequenced in this project, for which the same mutation affecting the equivalent residue in *h*IR has been identified (see below). The *tm1236* allele has an in-frame 561bp deletion in exon 14 that is predicted to remove 187 amino acids from the TK domain. The *m65* and *m631* alleles are both nonsense mutants of *daf-2* (Figure 5.5). The *m631* allele has an R281Stop substitution in the L1 domain and so is predicted to produce no functional protein. The *m65* allele has a W1449Stop substitution in the TK domain. It should be noted that Edward Campbell, a former research technician in the Gems lab, sequenced the *m65* allele.

5.3.3 Location of temperature-sensitive mutations across DAF-2 receptor:

The *daf-2* alleles for which sequence data exist are listed in Table 5.3.1 and Figure 5.1. The majority of class 1 alleles have mutations that fall into the α-subunit of the DAF-2 receptor. These include *m41*, which affects the CR domain. The alleles *m596, e1368* and *e1365/ sa193* affect the L2 domain, *sa229* is located in the boundary between the L2 and Fn0 domains. The *mu150* allele affects Fn0 while the *e1371* and *m212* alleles affect the α-subunit half of the Fn1 domain. The *m577* allele affects the β-subunit half of the Fn1 domain. The class 2 alleles are equally divided between the α-subunit, with *e979* affecting the L1 domain and *m579* and *sa187* both affecting the CR domain, and the TK domain of the β-subunit which contains the mutations caused by the *sa219, e1391* and *e1370* alleles.

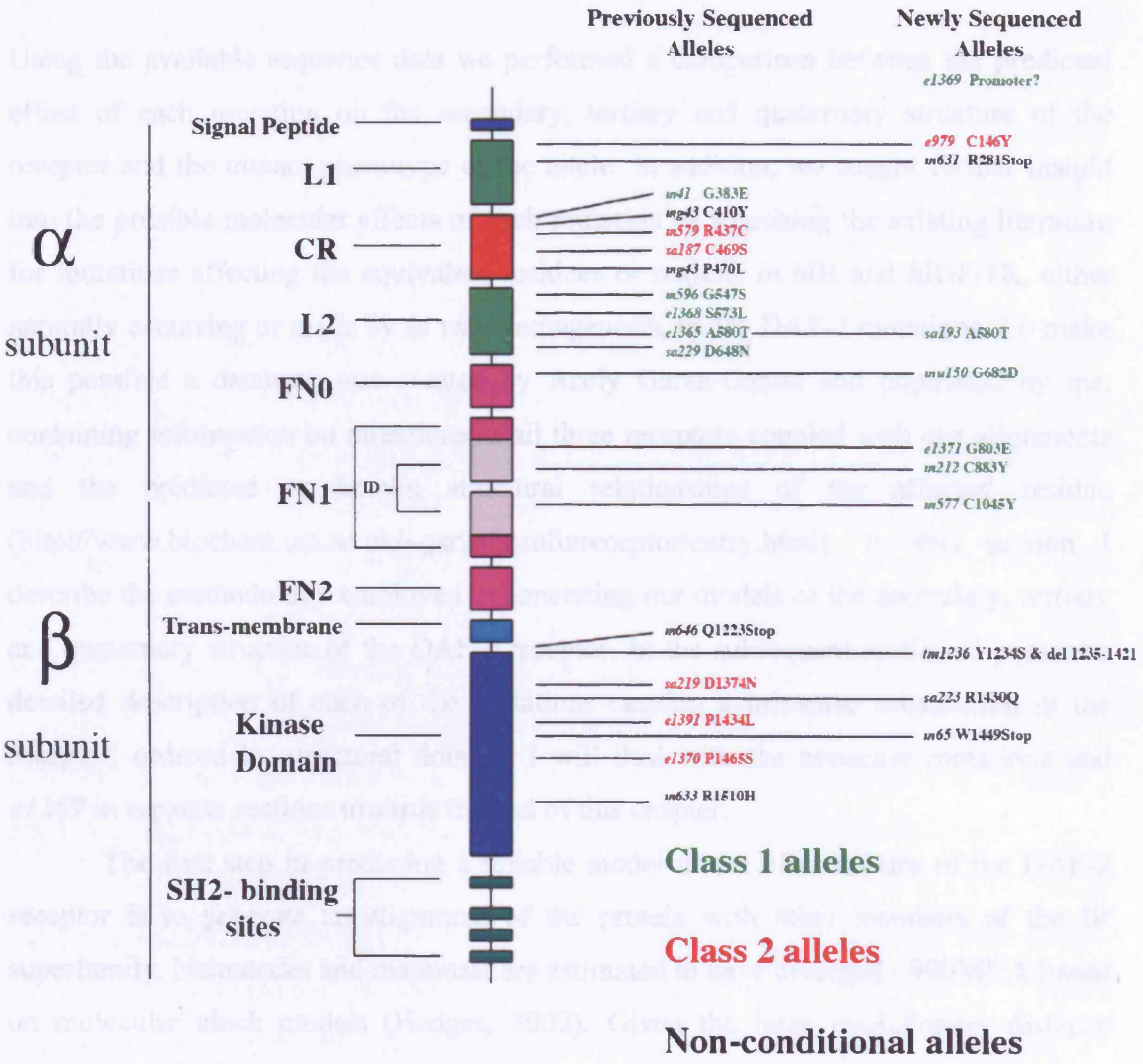
The location of the mutations shows only partial correlation with their phenotypic class. All class 1 mutations, except *e1369*, are located in the extracellular portion of the DAF-2 receptor. Class 2 alleles, however, are equally divided between the extracellular portion of the receptor and the cytoplasmic TK domain. Interestingly, none of the 23 alleles for which the sequence change has been identified affects either the N- or C-terminal extensions of the DAF-2 receptor.

Table 5.3.1 Mutations in *daf-2*

daf-2 allele	Class	Nucleotide change	Amino Acid Change	Domain	Equivalent *h*IR residue
e1369	**1**	**Not in CDS**	**None**	**-**	**-**
e979	**2**	**TGC-TAC**	**C146Y**	**L1**	**C8**
m631	**NC**	**CGA-TGA**	**R281STOP**	**L1**	**T113**
m41[b, c]	1	GGA-GAA	G383E	CR	Gap in alignment
mg43[a]	NC	TGT-TAT; CCC-CTC	C401Y; P470L	CR	C216; P289
m579[c]	2	CGT-TGT	R437C	CR	R252
sa187[a]	2	TGT-AGT	C469S	CR	C288
m596[c]	1	GGC-AGC	G547S	L2	G366
e1368[a]	1	TCA-TTA	S573L	L2	T392
e1365[a], ***sa193***	**1**	**TGC-TAC**	**A580T**	**L2**	**S399**
sa229[a]	1	GAT-AAT	D648N	L2/ Fn0	N470
mu150	**1**	**GGC-GAC**	**G682D**	**Fn0**	**G529**
e1371	**1**	**GGA-GAA**	**G803E**	**Fn1a**	**G623**
m212	**1**	**TGT-TAT**	**C883Y**	**Fn1a**	**C682**
m577*	**1**	**TGC-TAC**	**C1045Y**	**Fn1b**	**C798**
m646[d]	NC	CAA-TAA	Q1223STOP	JM	S967
sa219[a]	2	GAT-AAT	D1374N	TK	D1118
sa223	**NC**	**CGA-CAA**	**R1430Q**	**TK**	**R1174**
e1391[a]	2	CCC-CTC	P1434L	TK	P1178
m65	**NC**	**TGG-TAG**	**W1449STOP**	**TK**	**W1193**
e1370[a]	2	CCA-TCA	P1465S	TK	P1209
m633[d]	NC	CGT-CAT	R1510H	TK	R1253
tm1236	**NC**	**Deletion**	**Y1234S; $D^{1235-1421}$**	**TK**	**Y984; $D^{984-1165}$**

Mutations in DAF-2. The amino acid change, where known, is listed for each allele, as well as the nucleotide change. Alleles listed in **bold** were sequenced in this project. NC= non-conditional, CDS = coding sequence of gene. [a] Sequence data from Kimura *et al* (1997). [b] Data from Yu and Larsen (2001). [c] Data from Scott *et al* (2002). [d] K. Kimura and G. Ruvkun, personal communication. * The *e1365* allele characterised in Gems *et al* (1998) is actually the *m577* allele

Figure 5.1 Mutations in *daf-2*



Schematic representation of the DAF-2 proreceptor, labelled with the approximate location of all mutations that have been identified. Class 1 alleles are listed in green and class 2 alleles are listed in red. Non-conditional alleles are listed in black.

5.3.4 Integrated analysis of *daf-2* mutations

Using the available sequence data we performed a comparison between the predicted effect of each mutation on the secondary, tertiary and quaternary structure of the receptor and the mutant phenotype of the allele. In addition, we sought further insight into the possible molecular effects of each mutation by searching the existing literature for mutations affecting the equivalent residues or regions in *h*IR and *h*IGF-1R, either naturally occurring or made by *in vitro* mutagenesis, to our DAF-2 mutations. To make this possible a database was created by Acely Garza-Garcia and populated by me, containing information on mutations in all three receptors coupled with our alignments and the predicted or known structural relationships of the affected residue (http://www.biochem.ucl.ac.uk/~garza/insulinreceptor/entry.html). In this section I describe the methodology employed in generating our models of the secondary, tertiary and quaternary structure of the DAF-2 receptor. In the subsequent sections I present a detailed description of each of the mutations causing a missense substitution in the receptor, ordered by structural domain. I will deal with the nonsense mutations and *e1369* in separate sections towards the end of this chapter.

The first step in producing a reliable model of the 3D structure of the DAF-2 receptor is to generate an alignment of the protein with other members of the IR superfamily. Nematodes and mammals are estimated to have diverged ~990MYA based on molecular clock models (Hedges, 2002). Given the large evolutionary distance between the DAF-2 receptor and *h*IR and the fact that they share a relatively low level of sequence identity and similarity (see Table 4.3.6, in Chapter 4), we used a broad range of members of the IR superfamily for our alignments. These were from six different phyla, including the cnidaria, platyhelminthes, nematoda, mollusca, arthropoda and chordata.

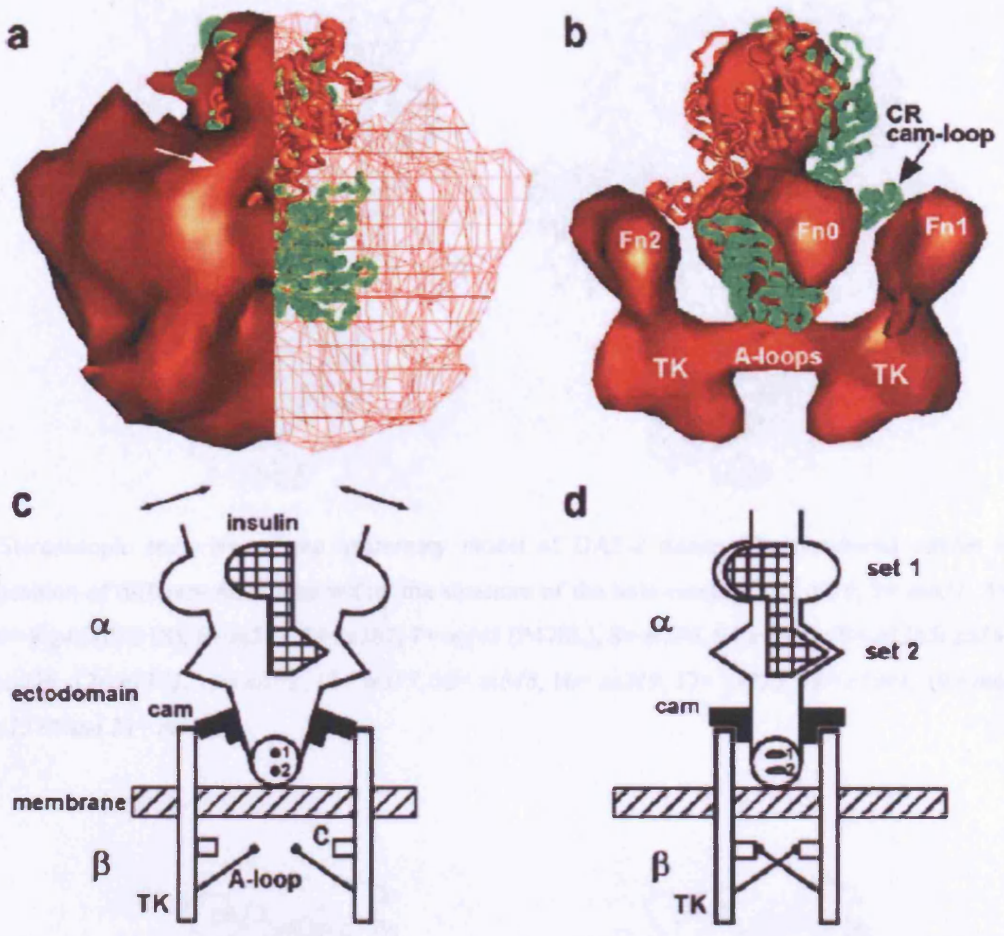
Crystal structures for homology modelling are available for the TK domains of both the *h*IR (Hubbard *et al.*, 1994) and *h*IGF-1R (Favelyukis *et al.*, 2001), as well as a fragment of the α-subunit of *h*IGF-1R (Garrett *et al.*, 1998), encoding the L1, CR and L2 domains. These were used as templates to model the corresponding domains of the DAF-2 receptor. Secondary and tertiary structure models of the DAF-2 were created by performing ClustalW alignments of our alignments (Thompson *et al.*, 1994). We used two different secondary structure prediction programs, JPred (Cuff *et al.*, 1998) and

PSIPRED (McAlduff *et al.*, 2002), to predict the location of α-helices and β-sheets in the DAF-2 protein. This information was used as a guide to our manual adjustments of the protein alignments. The manual adjustment of the protein alignments was carried out using the BioEdit sequence-editing program (see Chapter 4.2.6). Acely then took the adjusted alignments and threaded the wild-type DAF-2 amino acid sequence around the crystal structure template using two different programs, SWISS-MODEL (Schwede *et al.*, 2003) and MODELLER (Marti-Renom *et al.*, 2000) to generate the models of the L1-CR-L2 domains and the TK domains.

Unfortunately, there are no crystal structures available for any of the Fn domains from any member of the IR superfamily of proteins. Instead we made use of the many crystal structures of fibronectin type III domains from other proteins as templates to model the Fn domains of DAF-2. However, the degree of homology between these domains and any IR superfamily Fn domain is very low. For example there is only 23% similarity between *h*IR Fn0 and the closest fibronectin type III domain listed in the HOMSTRAD database (Mizuguchi *et al.*, 1998), which is the eighth fibronectin type III domain of the human fibronectin protein. Once the wild-type sequence models were constructed, Acely generated models for the mutations in DAF-2. We then compared the mutant and wild-type models for changes in secondary and tertiary structure.

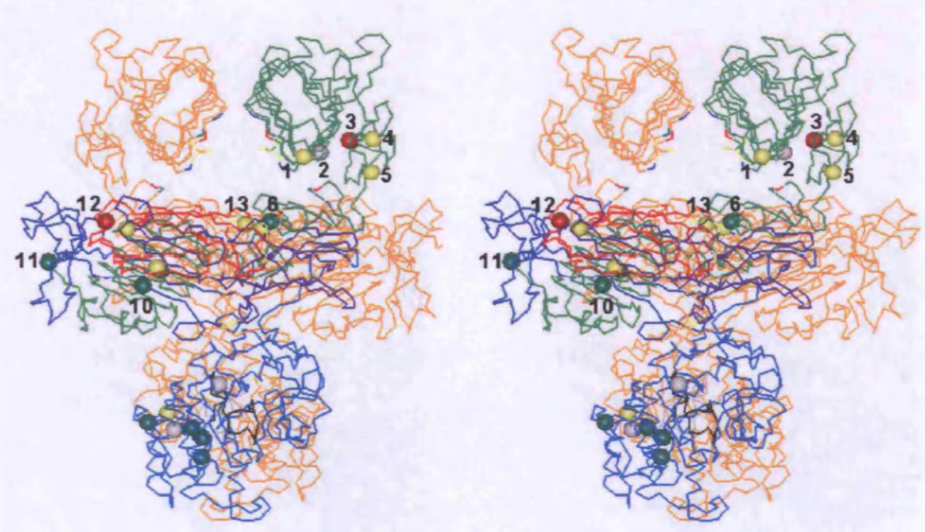
Prof. Peter Ottensmeyer, University of Toronto, carried out modelling of the quaternary structure of the DAF-2 receptor, using our alignments, based on a quaternary model of *h*IR produced by his group (Ottensmeyer *et al.*, 2000). This model provides a possible mechanism for signal transduction through *h*IR upon insulin binding. In this model the CR domains of each chain in the holo-receptor act as cams that hold the activation loops of each chain apart, which prevents autophosphorylation, in the absence of insulin (Figure 5.2c). However, the binding of insulin changes the conformation of the L1-CR-L2 region, such that the CR domains are moved upwards, allowing the activation loops to move closer and initiate transphosphorylation (Figure 5.2d). Upon dissociation of insulin the receptor reverts back to its unbound state, by the release of torsional energy from the strained disulphide bonds that link the chains. Unfortunately, the quaternary model of the DAF-2 receptor did not provide insight into the impact of the majority of our mutations on signal transduction through the receptor. However, this model may provide an insight into the effect of the *m212* mutation (see below).

Figure 5.2 Quaternary model of *h*IR and a possible signal transduction mechanism

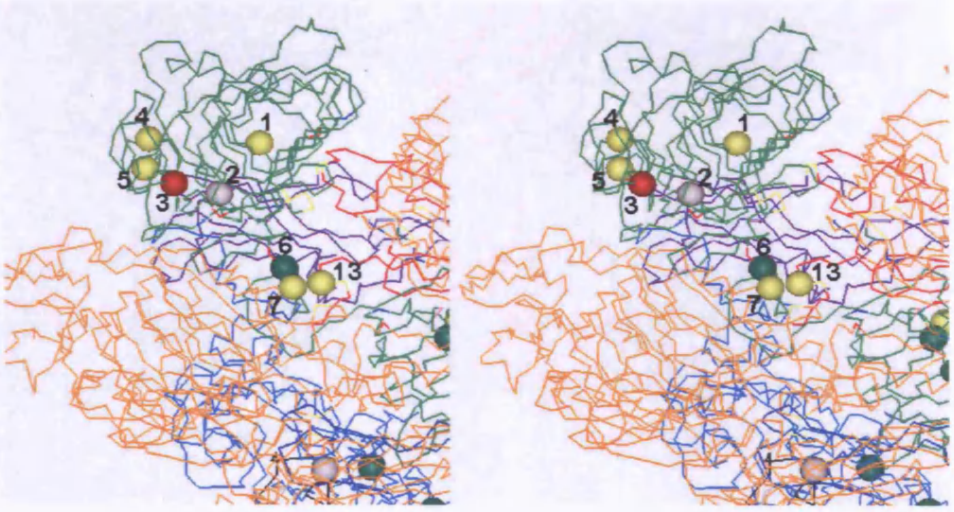


Panel **a**: End view of the full-mass representation of the IR dimer. Arrow: cam-like region on the CR domain. Panel **b**: Higher density solid surface representation of the view rotated slightly from panel a to show CR loop regions (cams) of atomic structure reaching the top portion of Fn2 domains. Panel **c**: Inhibitory state: ectodomain of dimeric R subunits each with two differing insulin binding sites and a blocking cam. Unbound bivalent insulin: α subunits resting against cams, crossing membrane, with TK domains separated. Abbreviations: A-loop, activation loop; C, catalytic region. Small on-axis circles (1 and 2) represent the α-α disulfide bonds. Arrows indicate thermally induced motion. Panel **d**: Insulin bound state: blocking cams rotated and α subunits resting closer to the center of the ectodomain. TK domains are in position for transphosphorylation via A-loops. Sets 1 and 2 indicate schematically different sets of amino acids from monomers I and II interacting with corresponding different sites on insulin. Image taken and legend adapted from (Ottensmeyer *et al.*, 2000).

Figure 5.3 Predicted quaternary structure of DAF-2 receptor

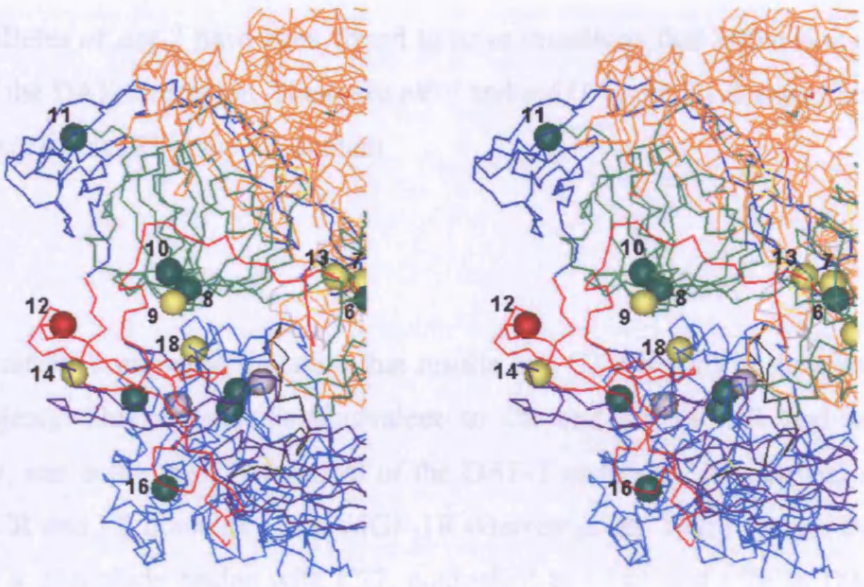


A: Stereoscopic end view of the quaternary model of DAF-2 dimer. The numbered circles are the position of different mutations within the structure of the holo-receptor. 1= *e979*, 2= *m631*, 3= *m41*, 4= *mg43* (C401S), 5= *m579*, 6= *sa187*, 7= *mg43* (P470L), 8= *m596*, 9= *e1368*, 10= *e1365/ sa193*, 11= *sa229*, 12= *e1371*, 13= *m212*, 14= *m577*, 15= *m646*, 16= *sa219*, 17= *sa223*, 18= *e1391*, 19= *m65*, 20= *e1370* and 21= *m633*.

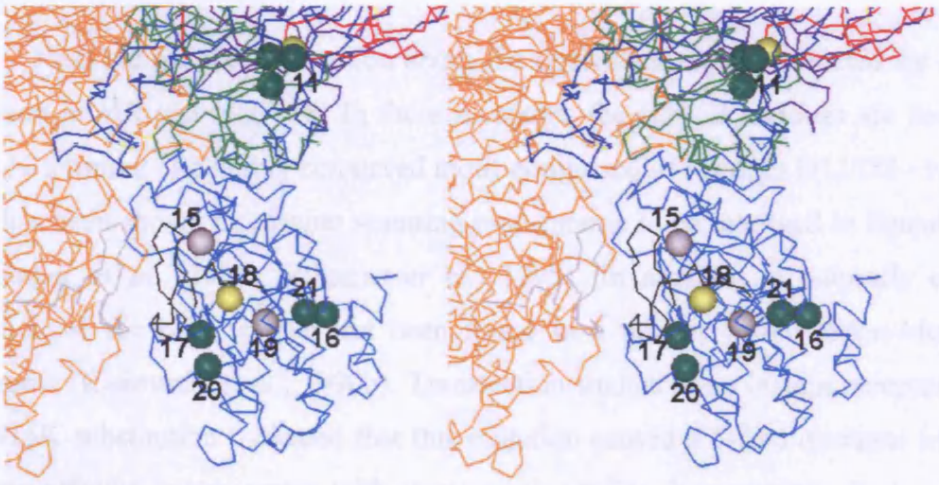


B: Close up of L1 and CR domains of DAF-2. The numbered circles are the position of different mutations within the structure of the holo-receptor. Numbers have same allele designation as above.

Figure 5.3 Predicted quaternary structure of DAF-2 receptor



C: Close up of L2 and Fn domains of DAF-2. The numbered circles are the position of different mutations within the structure of the holo-receptor. 1= *e979*, 2= *m631*, 3= *m41*, 4= *mg43* (C401S), 5= *m579*, 6= *sa187*, 7= *mg43* (P470L), 8= *m596*, 9= *e1368*, 10= *e1365/ sa193*, 11= *sa229*, 12= *e1371*, 13= *m212*, 14= *m577*, 15= *m646*, 16= *sa219*, 17= *sa223*, 18= *e1391*, 19= *m65*, 20= *e1370* and 21= *m633*.



D: Close up of TK domain of DAF-2. The numbered circles are the position of different mutations within the structure of the holo-receptor. Numbers have same allele designation as above.

5.3.5 Mutations affecting the receptor L-domains

In all, six alleles of *daf-2* have been found to have mutations that affect one of the L-domains of the DAF-2 receptor. These are *e979* and *m631* in the L1 domain, and *m596*, *e1368* and *e1365/ sa193* in the L2 domain.

e979

This allele carries a missense mutation that results in a C146Y substitution (sequenced in this project). This cysteine is equivalent to C8 and C3 of *h*IR and *h*IGF-1R, respectively, and is the most N-terminal of the DAF-2 mutations. The crystal structure of the L1, CR and L2 domains of the *h*IGF-1R (Garrett *et al.*, 1998) reveals that C3 is involved in a disulphide bridge with C22, equivalent to C181 and C26 in DAF-2 and *h*IR, respectively. The *e979* allele is class 2 and is the only allele that causes a 100% embryonic lethality/L1-stage arrest phenotype in its progeny at 25.5°C (Gems *et al.*, 1998; Vowels and Thomas, 1992). This allele also has a strong Daf-c phenotype at 15°C causing 20% dauer formation (Gems *et al.*, 1998). Potentially, the loss of this disulphide bond results in thermolabile protein that has reduced function at 15°C and no activity at 25°C.

Next we sought information about the equivalent regions affected by the *e979* mutation in *h*IR and *h*IGF-1R. In these receptors, the C8/ C3 residues are four amino acids N-terminal of a highly conserved motif comprised of residues D12/D8 - N15/N11. This has been shown by alanine scanning mutagenesis to be involved in ligand binding (Whittaker *et al.*, 2001; Williams *et al.*, 1995). In addition, a naturally occurring mutation of the N15 residue has been found in a patient with Rabson-Mendenhall syndrome (Kadowaki *et al.*, 1990b). Transfection studies of an insulin receptor bearing the N15K substitution indicated that this mutation caused a 5-fold decrease in insulin-binding affinity, in agreement with mutagenesis studies. Interestingly, it also caused a retardation of the transport of the mature receptor to the plasma membrane, leading to reduced levels of the receptor on the cell surface. In addition, these receptors also had a severely reduced ability to autophosphorylate upon insulin-stimulation. This suggests that the *e979* mutation may have type 2, 3 and 4 defects (see chapter 1.2.3 and Figure 1.4).

The C26 residue of *h*IR is N-terminal of V28 and G31, V183 and G186 in DAF-2, all three residues are part of the second repeat of a loosely conserved glycine-centred motif that appears at least four times in each L-domain (Bajaj *et al.*, 1987). The V28 and G31 residues are both affected by naturally occurring mutations that have been identified in two different patients with leprechaunism (van der Vorm *et al.*, 1992; Wertheimer *et al.*, 1994). In both cases the patients were heterozygotes for their respective mutations. Transfection studies of the V28A insulin receptor revealed that the mutation severely impairs proteolytic processing of proreceptor into the mature α- and β-subunits. This leads to a reduction in the number of receptors on the cell surface, the mutation had no effect on the insulin-binding affinity of the mature receptor or its ability to autophosphorylate in response to insulin (Wertheimer *et al.*, 1994). Similar transfection studies of the G31R mutant revealed a more severe phenotype than the V28A mutant. Here the proreceptor fails to be proteolytic cleaved and is not transported to the cell surface, the proreceptor was also unable to bind insulin or undergo autophosphorylation (van der Vorm *et al.*, 1992).

What can this tell us about the effects of *e979* on DAF-2 function? The *e979* mutation is predicted to remove a disulphide bond between C146 and C181 in the DAF-2 receptor. The *h*IR equivalents of these residues are C8 and C26. C8 is proximal to highly conserved residues involved in the post-translational processing of the mature receptor, and in ligand binding, moreover, the C26 residue is part of a glycine-centred motif that is critical in for correct proteolytic cleavage of the proreceptor. Taken together, this suggests that the *e979* mutation causes (1) severe impairment of post-translational processing of the DAF-2 receptor; (2) a reduction of the binding affinity of the receptor to its putative ligands: and/ or (3) its ability to autophosphorylate in response to ligand.

The C146Y substitution in *e979* is semi-conservative as both cysteine and tyrosine are hydrophobic residues, although tyrosine is less so due to the hydroxyl group present on its aromatic side-chain. Potentially at 15°C the C146Y mutation in *e979* has a less severe effect on the post-translational processing of the DAF-2 receptor due to semi-conservative nature of the substitution and so phenotypically resembles the N15K and V28A mutations of *h*IR. At 25.5°C the loss of the disulphide bridge may completely inhibit the proteolytic processing of the DAF-2 proreceptor in a similar way to the G31R mutation in *h*IR resulting in a complete lack to mature DAF-2 receptors at

the cell surface, which in turn leads to the embryonic lethality/ L1-stage arrest of progeny phenotype at this temperature.

m596

This allele of *daf-2* was originally classified as a class 2 allele based on the fact that it develops Unc and Eat phenotypes at 25.5°C (Gems *et al.*, 1998). However, it is suppressed by *daf-12(m20)* (Gems *et al.*, 1998) and is only weakly Hyp (hypoxia resistant) (Scott *et al.*, 2002) and so has been reclassified as a class 1 allele. *m596* has a missense mutation causing a G547S substitution in the L2 domain of the DAF-2 receptor (Scott *et al.*, 2002), equivalent to G366/ G356 in *h*IR and *h*IGF-1R, respectively. According to our model of the DAF-2 receptor the G547 residue is buried in the structure of the L2 domain and located in a turn between the fifth and sixth β-strands. Mutation of this residue to serine is expected to cause clashes with several of the neighbouring buried residues, especially N518/ N329/ S339 in DAF-2, *h*IGF-1R and *h*IR, respectively, disrupting the stability of the fold.

G547 in DAF-2, like G31 in *h*IR, is the central glycine in the second repeat of the four glycine-centred motifs found in each L-domain (Bajaj *et al.*, 1987). Like the above-mentioned G31R mutation in *h*IR identified in a heterozygote patient with leprechaunism (van der Vorm *et al.*, 1992), a G366R mutation has also been identified in a patient with leprechaunism (Wertheimer *et al.*, 1994). Transfection studies of *h*IR carrying the G366R substitution have shown that this mutation impairs the post-translational processing of the receptor leading to a reduced number of mature receptors on the cell surface. Unlike the G31R mutation (van der Vorm *et al.*, 1992), the G366R mutation seems to have no effect on the insulin binding affinity or the autophosphorylation of the receptor (Wertheimer *et al.*, 1994).

We postulate, on the basis of the above, that this disruption would also lead to a similar impairment of the post-translational processing of the DAF-2 receptor, like the G366R mutation in *h*IR, reducing the number of mature receptors on the cell surface without effecting either DAF-2 ligand binding or its autophosphorylation.

e1368 and *e1365/ sa193*

Three further alleles of *daf-2* mapping to the L2 domain are *e1368,* which causes a S573L substitution (Kimura *et al.*, 1997), and *e1365* and *sa193*, which have an identical nucleotide change (see table 5.3.1) leading to an A580T substitution (Kimura *et al.*, 1997 and this study). According to the model of the DAF-2 receptor, both residues are buried within the structure of the L2 domain, so their mutation is likely to destabilise the fold.

No mutations have been found at equivalent positions in any member of the IR superfamily. However, these residues are part of the fourth repeat of the glycine-centred motif in the L2 domain (Bajaj *et al.*, 1987). In the wild-type DAF-2 sequence the nearby A571 residue occupies the central role played by G390 in the equivalent fourth repeat in the L2 domain of *h*IR. In all other nematode DAF-2 proteins this glycine is conserved. The proximity of the S573L mutation in the *e1368* allele to the A571 residue suggests that it might interfere with the role of the alanine residue in the turn between the eighth and ninth β-strands of the L2 domain. This could result in an impairment of the post-translational processing of the DAF-2 receptor, as does the G547L mutation in the *m596* allele. However, the effect must be weaker since *e1368* is a class 1A allele where dauers begin to recover after 24 hours at 25.5°C, unlike *m596* dauers (Gems *et al.*, 1998).

e1365/ sa193

The *e1365/ sa193* mutation A580T (*e1365* sequenced by Kimura *et al*, 1997 and *sa193* in this project) affects an alanine residue that is conserved in all nematode DAF-2 receptors and two arthropod proteins, DInR and BIR. However, this alanine is replaced with a serine in most vertebrate members of IR superfamily (except *h*IRR where it is replaced with a threonine). The equivalent residue in the fourth repeat of this motif in the L1 domain is also an alanine residue in almost all IR superfamily members, A261/ A92 in DAF-2 and *h*IR, which suggests some functional significance. Given that the *e1365/ sa193* mutation affects the same repeat as the *e1368* mutation, suggests that it might cause an impairment of the post-translational processing of the DAF-2 receptor. A similar type of defect would be consistent with the similar phenotype of *e1368* and *e1365/ sa193*, both are class 1A alleles (Gems *et al.*, 1998).

5.3.6 Mutations in the cysteine-rich (CR) domain

The CR domain of most IR superfamily members is made up of eight irregular modules, the first three of which contain four cysteine residues, C1-C4, linked in the pattern C1:C3 and C2:C4. The other five modules contain only a single pair of cysteines that are bonded to each other. Four alleles, *m41*, *mg43*, *m579* and *sa187*, have mutations that affect residues in the CR domain (Table 5.3.1).

m41

This allele is the only class 1 allele of *daf-2* with a missense mutation in the CR domain. The *m579* and *sa187* alleles are both class 2, while *mg43* is a non-conditional allele and has two different mutations within the CR domain. The *m41* allele causes a G383E mutation in module 2 of the CR domain (Yu and Larsen, 2001)(D. McCulloch and D. Gems, unpublished data). This module is unusual in all nematode DAF-2 proteins in that it lacks the C2 and C4 cysteines, equivalent to C189 and C200 in the *h*IGF-1R structure. We postulate that in nematodes this disulphide bond is replaced by an intradomain salt bridge, in the case of the *C. elegans* DAF-2 between E381 and R392. While the G383 residue is conserved amongst all nematodes, it falls into a gap in the alignment with all other members of the IR superfamily, making it difficult to predict the effect of its mutation on the structure of DAF-2.

mg43

This non-conditional allele has two different mutations affecting the CR domain (Kimura *et al.*, 1997). The first is a C401Y substitution in module 3 of the CR domain, equivalent to C216/ C209 in *h*IR/ *h*IGF-1R. This is the third cysteine in module three and predicted to form a disulphide bond with C393 in DAF-2, C208/ C201 in *h*IR/ *h*IGF-1R. The second mutation in the *mg43* allele is a P470L substitution of an entirely conserved proline in module seven of the CR domain. Although no mutations affecting either of these residues have been found in any other IR superfamily member, mutations of similar residues have been noted in *h*IR. A H209R substitution in *h*IR, adjacent to the C208 residue that is bonded to C216, has been shown to impair several different steps in

the post-translational processing of the proreceptor so that only 10% of the mature receptor is transported to the cell surface. These receptors retain wild-type binding affinity and normal tyrosine kinase activity (Kadowaki *et al.*, 1991). Given that the C401 residue in DAF-2 is bonded to C393, which is equivalent to C208 in *h*IR, it is possible that the substitution of this residue to tyrosine in the *mg43* allele also causes an impairment of the post-translational processing of the DAF-2 proreceptor.

The second mutation in the *mg43* is a P470L substitution in module seven of the CR domain. A similar proline to leucine substitution is seen in the first module of the CR domain of *h*IR at position 193 (Carrera *et al.*, 1993). Like P470, the P193 residue is adjacent to a cysteine involved in a disulphide bridge, the P193L mutation in *h*IR has been shown to prevent dimerisation of the proreceptor, inhibit cleavage and block transport to the cell surface (Maggi *et al.*, 1999), although some mature receptor does get to the cell surface and appears to have normal insulin binding affinity (Takata *et al.*, 1994) and undergoes normal autophosphorylation (Carrera *et al.*, 1993). Potentially, both mutations in the *mg43* allele cause defects in the post-translational processing of the DAF-2 proreceptor, resulting in the non-conditional phenotype of this allele. It is possible that each mutation on its own would display a class 2 phenotype.

sa187

The idea above is supported by the mutation in *sa187*, a class 2 allele (Gems *et al.*, 1998). The *sa187* allele has a C469S substitution in module 7 of the CR domain (Kimura *et al.*, 1997). This residue is adjacent to the P470 residue mutated in *mg43* and is predicted to form a disulphide bond with C483. Given the fact that it is adjacent to P470 it seems likely that the C469S mutation also leads to the same post-translational defects as the P470L mutation.

m579

This is a class 2 allele and has a R437C substitution (Scott *et al.*, 2002) in the fifth module of the CR domain. The mutation in *m579* is interesting as it is the only mutation in the extracellular domain of the DAF-2 receptor for which an identical mutation in *h*IR has been identified. A patient with type A insulin resistance accompanied by acanthosis nigricans was found to be homozygous for a missense mutation causing a

R252C substitution in the CR domain of *h*IR (Hamer *et al.*, 2002). Erythrocytes cultured from the patients revealed a 40% decrease in receptor levels on the cell surface and a five-fold decrease in insulin binding affinity when compared to cell from a healthy individual.

Transfection studies of the R252C mutant in CHO cells revealed that the receptor has a delayed maturation of the proreceptor into the α- and β-subunits, and reduced transport of the mature receptor to the cell surface. The mutant receptor also had reduced insulin binding affinity and an almost complete lack of insulin-stimulated receptor internalisation. Interestingly the R252C mutant receptor seemed to undergo normal autophosphorylation and retained a normal ability to phosphorylate the IRS-1 and -2 adaptor proteins as well as Akt, part of the PI3 kinase pathway. However the mutant receptor had marked reduction in its ability to induce Shc and ERK1/2 phosphorylation, as well as reduced incorporation of thymidine into DNA. This difference in substrate phosphorylation may be due to prolonged exposure of the mutant receptor to insulin due to the reduced internalisation of the receptor, suggesting that receptor-mediated endocytosis may act as an important attenuator of insulin signalling.

Hamer *et al* (2002) suggest that the R252C substitution may cause an aberrant disulphide bond with C3 in the L1 domain, locking the two domains into a conformation that is unfavourable to high-affinity ligand binding. This is likely to be incorrect, since two patients with a R252H substitution in *h*IR also suffered from type A insulin resistance and the mutant receptors also have reduced receptor levels on the cell surface as well as reduced insulin binding affinity (Maassen *et al.*, 2003; Nakashima *et al.*, 1995). We propose instead that the molecular defects in the *m579* allele, as well as the R252C and R252H *h*IR mutants is due to the loss of the arginine residue. This residue is conserved in nearly all members of the IR superfamily and according to our model of the DAF-2 receptor R437 forms two hydrogen bonds with E185, E30/ E26 in *h*IR and *h*IGF-1R, in the L1 domain (see Figure 5.4). This interaction is crucial for keeping the L1 and CR domain in their correct orientation and its loss would undoubtedly affect ligand binding in these receptors.

Figure 5.4 Hydrogen bonds between R437 and E185 in wild-type DAF-2

Cartoon of the structure of the L1-CR-L2 domains of the DAF-2 receptor. In the wild-type receptor the E185 residue in the L1 domain is predicted to form two hydrogen bonds with the R437 residue in the CR domain, based on the *h*IGF-1R L1-CR-L2 crystal structure (Garrett *et al.*, 1998). In the *daf-2(m579)* allele the R437 residue is mutated to a cysteine that can no longer take part in these hydrogen bonds, leading to a change in the orientation of the L1 domain with respect to the L2 domain.

5.3.7 Mutations in the fibronectin type III (Fn) domains

Five *daf-2* alleles, all of which are class 1, have mutations affecting the Fn domains of the DAF-2 receptor. The *sa229* allele actually falls into the boundary between the L2 and Fn0 domains, while the *mu150* allele actually affects the Fn0 domain. The *e1371* and *m212* alleles affect the α- subunit half of the Fn1 domain while the *m577* allele affects the β-subunit half of the Fn1 domain.

sa229

This allele causes a D648N substitution (Kimura *et al.*, 1997). Given its position at the boundary between the L2 and Fn0 domains it is impossible to predict the structural impact of this mutation. Moreover, the D648 residue, N470/ S460 in *h*IR and *h*IGF-1R, is poorly conserved across IR superfamily members. However, in the *h*IR two different mutations at the C-terminus of the L2 domain, K460E and N462S are proximal to the boundary between L2 and Fn0. Both increase the rate of receptor internalisation and degradation leading to a reduced number of receptors on the cell surface (Kadowaki *et al.*, 1990a; Kadowaki *et al.*, 1990c). Possibly, *sa229* has a similar effect.

mu150

This allele has a G682D (G502 in *h*IR) substitution in the C strand of the Fn0 domain (sequenced in this project). No mutation of the equivalent residue has been found in any other member of the IR superfamily. However, a mutation causing an I485T substitution in the B strand of the Fn0 domain in *h*IR was identified in a patient with type A insulin resistance and lipodystrophy (Yokota *et al.*, 1990). This mutation did not cause the receptor to have reduced insulin-binding affinity when expressed in CHO cells. But the patients own cells did have a reduced level of insulin binding suggesting that the mutation caused a reduction in the receptor level on the cell surface. Potentially, a similar effect is caused by *mu150* allele. However, the impairment is likely to be weaker as the *mu150* allele is class 1A and dauers formed at 25.5°C recover within 24 hours (Garigan *et al.*, 2002).

e1371

This allele has a G803E substitution in the α- subunit half of the Fn1 domain (sequenced in this project). This glycine residue, G623/ G609 in *h*IR and *h*IGF-1R, is part of a motif P-X-X-PNG that is highly conserved in our alignment and is located in between the B and C strands of the Fn1 domain. Interestingly, the same motif also appears between the B and C strands of the Fn2 domain. This motif is not a generally conserved feature of the fibronectin type III structures listed in the HOMSTRAD database. A PROSITE scan (Falquet *et al.*, 2002) of the Protein Data Bank for the P-X_2-PNG motif retrieved more than 100 entries. This motif is generally found in tight turns, mostly type βI, in which the glycine residue displays positive φ and ψ angles. This would be an unfavourable conformation for the substituted glutamic acid residue for the mutation in the *e1371* allele.

m212

This allele is an atypical class 1 allele, i.e. it has a strong Daf-c phenotype at all temperatures, but is fully suppressed by *daf-12(m20)* (Gems *et al.*, 1998). *m212* allele has a C883Y substitution (sequenced in this project), C682/ C669 in *h*IR and *h*IGF-1R, and is located in the large insert between the C and C' strands of the Fn1 domain. In *h*IR the α- and β-subunits are linked by a single disulphide bond between C647 and C872 to form an α-β dimer (Cheatham and Kahn, 1992; Sparrow *et al.*, 1997). These α-β dimers are then linked by four class 1 disulphide bridges between the adjacent α-subunits to form the mature tetrameric $\alpha_2\beta_2$ receptor. The four cysteines involved in the α-α subunit disulphide bonds are C524 (Bilan and Yip, 1994; Macaulay *et al.*, 1994; Schaffer and Ljungqvist, 1992), C682, C683 and C685 (Lu and Guidotti, 1996; Sparrow *et al.*, 1997; Wu and Guidotti, 2002; Wu and Guidotti, 2004), equivalent to C706, C883, C884 and C886 in the DAF-2 receptor. Site-directed mutagenesis of all four of these cysteines to serines in *h*IR results in free α-β dimers that are incapable of autophosphorylation and only transported to the cell surface at 50% of wild-type levels, although at sufficiently high concentration in the plasma membrane these dimers can associate non-covalently to produce functional tetrameric receptors (Wu and Guidotti, 2002).

When C682 is substituted to serine the majority of protein expressed in COS7 cells is present in the mature tetrameric form, however a small proportion is present as free α-β dimers (Lu and Guidotti, 1996). When C524 is substituted with tryptophan the majority of the protein expressed in COS7 cells is not transported to the cell surface, so that levels of tetrameric mature receptor are only 12% of wild-type *h*IR controls (Wu and Guidotti, 2004). Wu and Guidotti (2004) suggest that the bulky tryptophan residue causes misfolding of the proreceptor, leading to its retention in the endoplasmic reticulum and subsequent degradation. A similar effect may be occurring with the C883Y mutation in the *m212* allele. The tyrosine residue has a bulky aromatic side-chain that may also cause misfolding of the DAF-2 proreceptor leading to reduced levels of the mature receptor on the cell surface. Also this mutation may also cause a small proportion of the DAF-2 protein to remain as free α-β dimers on the cell surface, which would be incapable of ligand-stimulated autophosphorylation as seen for the C682S mutation (Lu and Guidotti, 1996). This could explain the severity of the Daf-c phenotype of this allele.

In the quaternary model of the *h*IR presented by Ottensmeyer *et al* (2000), the disulphide bonds involved in the α-α subunit interaction are crucial components of the signal transduction mechanism of the receptor. In this model, the cysteine residues provide torsional resistance that maintains the 'cams' formed by the CR domains in the position that blocks transphosphorylation. Upon insulin binding, the bonds between these residues become strained, and the 'cams' are moved upwards so that the activation loop of each kinase domain is able contact the other and initiate autophosphorylation. When the insulin molecule dissociates from the receptor, the torsional energy stored by the disulphide bonds is released and this returns the receptor to the inactive conformation. Ottensmeyer *et al* (2000) point out in their model that loss of one these disulphide bonds may result in the loss of torsional resistance. This in turn would make the approach of the activation loops of each kinase domain more random, and so less efficient at transphosphorylation. This result seems to correlate well with the biochemical data above.

m577

This allele has a mutation in the F strand of the β-subunit half of the Fn1 domain, causing a C1045Y substitution (sequenced in this project). The C1045 residue is expected to form a disulphide bond with C1055 based on homology with C798 and C807, the equivalent residues in *h*IR, which have been shown to be linked by a disulphide bridge (Sparrow *et al.*, 1997). Site-directed mutagenesis of C798 in *h*IR and C776 in *h*IGF-1R to serines impairs the post-translational processing of the proreceptors of both proteins leading to a complete lack of mature receptors on the cell surface (Maggi and Cordera, 2001). However, site-directed mutagenesis of the C807 residue to serine does not seem to have any affect on the post-translational processing of the *h*IR proreceptor or the insulin binding affinity of the mature receptor or its ability to phosphorylate IRS-1 (Maggi *et al.*, 1998). Given that the C798 residue has a disulphide bond to C807 (Sparrow *et al.*, 1997), it is hard to reconcile the very different effects of mutating each cysteine residue to a serine. It is unlikely that C1045Y mutation in the *m577* allele would lead to a complete lack of mature DAF-2 receptor on the cell surface, as this allele is class 1 and forms dauers only at 25°C (Gems *et al.*, 1998). However, the concept of the C1045Y mutation leading to a receptor with normal ligand binding and substrate phosphorylation but reduced levels on the cell surface seems to be consistent with the other class 1 alleles of *daf-2*.

5.3.8 Mutations in the TK domain

Seven alleles of *daf-2* have mutations that affect the TK domain of the DAF-2 receptor. These are *sa219*, *sa223*, *e1391*, *m65*, *e1370*, *m633* and *tm1236*. An additional allele *m646* has a mutation that affects the juxtamembrane region that is N-terminal of the TK domain. All of the mutations that affect the cytoplasmic portion of the β-subunit are either class 2 or non-conditional alleles.

sa219

This allele causes a D1374N substitution in the αE helix of the TK domain (Kimura *et al.*, 1997). This residue is highly conserved across our alignment and the equivalent

residues D1118 and D1091 in the *h*IR and *h*IGF-1R crystal structures (Favelyukis *et al.*, 2001; Hubbard *et al.*, 1994) are buried, and their carboxyl Oδ atoms act as acceptors in three hydrogen bonds. One oxygen atom interacts simultaneously with the protonated Nδ1 atom the of H1058/ H1031 residues in *h*IR and *h*IGF-1R (F1310 in DAF-2) located in the loop that spans αC and β4 regions of the TK domain, and with the NH-group of F1276/ F1249/ F1533 residue in each receptor situated after the αJ α–helix. The other oxygen atom forms a hydrogen bond with the NH-group of F1277/ Y1250/ V1534 residue of each receptor. These hydrogen bonds play an important role in keeping the αJ helix in the correct orientation. The asparginine residue substituted for the aspartic acid in *sa219* lacks one of the Oδ atom and so is not able to establish the correct arrangement of hydrogen bonds. This may destabilise the fold, potentially affecting the activity of the kinase domain.

sa223

sa223 is an unusual non-conditional allele with a mutation causing a R1430Q substitution in the P+1 loop of the TK domain of the DAF-2 receptor (sequenced in this project). This loop is important in determining the substrate specificity of the kinase domain for either tyrosine residues or serine/ threonine residues (Hubbard *et al.*, 1994). The R1430 residue is equivalent to R1174 in *h*IR and the same arginine to glutamine substitution of this residue has been reported in several heterozygous individuals with severe type A insulin resistance (Hojlund *et al.*, 2004; Krook *et al.*, 1995; Moller *et al.*, 1991; Moritz *et al.*, 1994). Transfection studies of the R1174Q mutant receptor in CHO cells suggest that the basal level of receptor autophosphorylation is comparable to wild-type receptors. However, insulin-stimulated autophosphorylation is only 30% of wild-type controls (Krook *et al.*, 1996; Moritz *et al.*, 1994). At high concentrations of insulin the mutant receptors were able to phosphorylate the IRS-1 adaptor protein to ~60% of wild-type levels, but were not able to induce the insulin-stimulated incorporation of thymidine into DNA or phosphorylate the MAP kinase protein (Krook *et al.*, 1996; Krook *et al.*, 1997b).

In the structures of *h*IR and *h*IGF-1R kinase domains, both Nη atoms of the R1174/ R1147 residue are hydrogen-bonded to the backbone oxygen atoms of the P1209/ P1182 and L1213/ L1186 residues. The loss of one of these interactions

probably destabilises the conformation of the P+1 loop. Both of these residues are also conserved in the DAF-2 receptor, P1465 and L1469, respectively. It therefore seems likely that the R1430Q mutation in *sa223* has a similar effect on the kinase activity of the DAF-2 receptor to the R1174Q mutation in *h*IR.

e1370

Interestingly, the P1465 residue that R1430 is expected to form a hydrogen bond with in the DAF-2 receptor is the site of the mutation in the *e1370* allele (Kimura *et al.*, 1997). This canonical *daf-2* allele is class 2 and has a proline to serine substitution; no human equivalent of this mutation has been reported. This residue is part of a hydrophobic core and the replacement of the hydrophobic proline with the polar serine is likely to destabilise the fold and reduce the kinase activity of the TK domain.

e1391

This allele is class 2 and has a P1434L substitution (Kimura *et al.*, 1997), which also affects the P+1 loop of the DAF-2 receptor. This proline residue is also conserved in *h*IR, P1178, and the same substitution for leucine has been identified in heterozygous patient with severe type A insulin resistance (Krook *et al.*, 1994). Transfection studies of P1178L mutant receptors in CHO cells revealed them to have similar properties to the R1174Q mutant receptors (Krook *et al.*, 1996; Krook *et al.*, 1997b). However the P1178L mutant receptors could phosphorylate the IRS-1 adaptor protein to ~80% of wild-type levels, when stimulated with high insulin levels (Krook *et al.*, 1997b). The proline to leucine substitution is expected to cause unfavourable steric clashes within the structure of the kinase domain, which may explain the slightly weaker impairment of the kinase activity of the P1178L mutant compared to the R1174Q mutant receptors. Potentially this is why the *e1391* allele displays class 2 phenotypes but *sa223* is non-conditional.

m633

This non-conditional allele has a R1510H substitution (K. Kimura and G. Ruvkun, personal communication) affecting the loop between the αH and αI helices of the DAF-

2 TK domain. This residue is conserved across virtually all kinases and is equivalent to R1253 in *h*IR. This residue is structurally important for the stability of the fold as it forms a salt bridge with E1179 that is a key interaction to bring together α-helix αEF and the loop that spans α-helixes α-H and α-I. No equivalent mutations have yet been reported in *h*IR, but it is likely that in DAF-2 this mutation severely impairs the activity of the kinase domain, causing its non-conditional phenotype.

5.3.9 *e1369* an allele with a mutation outside the *daf-2* open reading frame

This allele is an atypical class 1 allele and has the most severe Daf-c phenotype of the ts alleles characterised in Gems *et al* (1998), yet is fully suppressed by *daf-12(m20)*. Both strands of cDNA, synthesised from two independent RNA extractions of *daf-2(e1369)* worms, were sequenced and no mutations affecting the cDNA were detected. In order to ensure that the *e1369* allele used in the sequencing, which phenotyped as in Gems *et al* (1998), was actually a *daf-2* mutant I performed complementation tests with two different alleles of *daf-2* for which the mutations had been identified. The *e1369* allele failed to complement either *daf-2(m41)* or *daf-2(m577)* indicating that it is an allele of *daf-2* (data not shown).

Given that the full-length cDNA from *e1369* has no mutation, the DAF-2 receptor produced by this allele must have a wild-type amino acid sequence. The *E19* allele of *dinr* also has no mutation affecting the coding sequence of the gene (Tatar *et al.*, 2001). Interestingly, this is the only known allele of the *dinr* gene that is homozygous viable. The lack of a mutation in the coding sequence suggests that both *daf-2(e1369)* and $dinr^{E19}$ may have mutations in regulatory regions of their respective genes that cause reduced transcription of mRNA.

There are several reported cases of heterozygous patients with either leprechaunism or severe type A insulin resistance in which one of the alleles has been shown to have reduced levels of transcription of the *INSR* gene, whilst having no mutation in the coding sequence of that allele (Imano *et al.*, 1991; Kadowaki *et al.*, 1990d; Longo *et al.*, 1992; Ojamaa *et al.*, 1988; Suzuki *et al.*, 1995). There is also a reported case of a patient with severe type A insulin resistance with acanthosis nigricans who appears to be a compound heterozygote for two different mutations that cause a

decrease in the level of transcription of the *INSR* gene (Imano *et al.*, 1991), although the exact location of the mutation has not been determined in any of these cases.

In an attempt to locate the mutation in the *e1369* allele ~5.2Kb of genomic DNA upstream of the trans-splice acceptor of exon 1 and the whole of intron 1 (~5.5Kb) of the *daf-2* gene were sequenced. No mutations in either of these regions were detected. Therefore any mutation in the *e1369* allele must lie either further upstream than 5220bp of the trans-splice acceptor of exon 1 or potentially in one of the other introns of the *daf-2* gene.

5.3.10 Nonsense mutations and the deletion allele of *daf-2*

In this section I deal with three nonsense alleles (*m65*, *m631* and *m646*) and a deletion allele (*tm1236*) of *daf-2*. All of these alleles are non-conditional. Epistasis analysis on the three nonsense alleles using null alleles of *daf-16* and *daf-18* has revealed that they can be placed into an allelic series based on their suppression by the null alleles above (M. Nanji and D. Gems, unpublished results). In terms of their DAF-2 activity the series is as follows: *m65*< *m631*< *m646*. Below I detail the mutation in each of the alleles, in the order of their premature stop codon, and the possible effect on the receptor.

m631

This allele has an opal mutation at codon 281 in the L1 domain (sequenced in this project) of the DAF-2 receptor sequence replacing an arginine residue with a stop (see Figure 5.5). This residue is equivalent to T113/ T107 in *h*IR and *h*IGF-1R. Of the three nonsense alleles this mutation is predicted to truncate the protein closest to the N-terminus of the receptor resulting in no functional protein. However, based on epistasis data on the suppression of the Daf-c phenotype, we postulate that *m631* is not a molecular null mutant (M. Nanji and D. Gems, unpublished results) and that there is readthrough of this stop codon to produce some functional DAF-2 protein.

Nonsense mutations in the human insulin receptor gene, leading to premature termination of translation, are frequently associated with decreased levels of mRNA (Taylor *et al.*, 1992). Recently, a compound heterozygote patient was identified carrying the mutations R108Q and K115N of the L1 domain of the *h*IGF-1R, N282 and D289 in

DAF-2. This resulted in a third of the normal IGF-1 binding affinity leading to growth retardation (Abuzzahab *et al.*, 2003). This suggests that this region is involved in ligand binding, thus, any functional DAF-2 protein produced by readthrough in *m631* allele would be predicted to have reduced binding affinity for DAF-2 ligands, as well as having reduced levels of DAF-2 receptor due to decreased mRNA levels caused by the premature stop codon.

m646

This allele has an ochre mutation that affects residue Q1223 located in the juxtamembrane region of the DAF-2 receptor (K. Kimura and G. Ruvkun, personal communication). Like the *m631* allele, we do not believe that the *m646* mutation represents a null allele of *daf-2*. This allele is predicted to have greater DAF-2 activity than the *m631* and *m65* alleles (M. Nanji and D. Gems, unpublished results). The epistasis results suggest that some functional DAF-2 protein must be produced by readthrough of the ochre stop codon. The severity of the Daf-c phenotype of this allele may be due to the combined effects of reduced protein level due to the nonsense mutation and a decrease in the levels of *daf-2* mRNA, also caused by the premature stop codon.

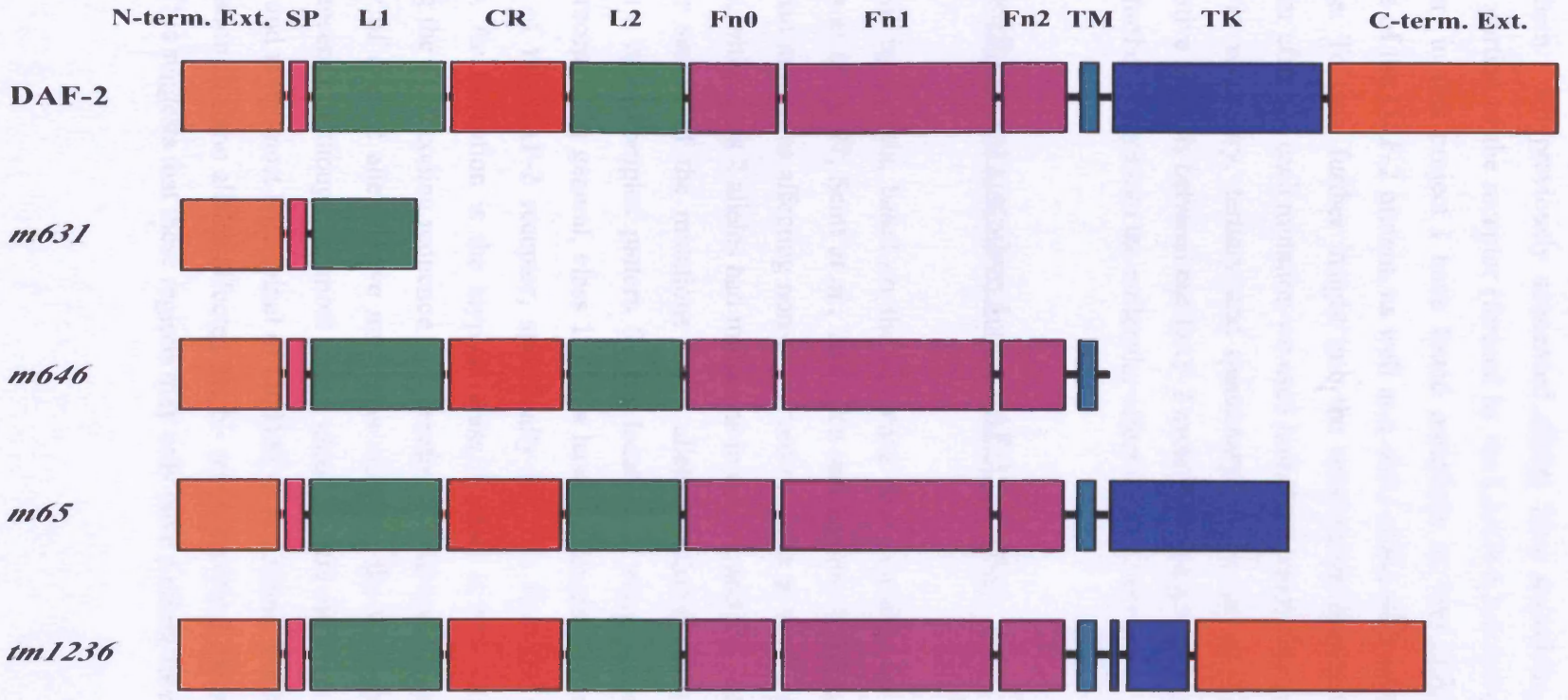
m65

This allele has an amber mutation affecting codon 1449 (sequenced in this project), encoding a tryptophan residue, of the DAF-2 receptor. This residue is equivalent to W1193 in the αF helix of the *h*IR TK domain and is completely conserved across our alignment. Substitution of W1193 to leucine has been reported for two unrelated families (Imamura *et al.*, 1994). Transfection studies of the W1193L mutant receptor in COS7 and Rat-1 cells indicated that insulin-binding affinity was normal. However, the insulin-stimulated autophosphorylation of the mutant receptors was ~80% of wild-type controls. In addition, the rate of proreceptor degradation was found to be 2-fold higher in this mutation than wild-type controls leading to a reduced number of mature receptors on the cell surface (Imamura *et al.*, 1994). At present *m65* represents the best candidate for a null allele of *daf-2*.

tm1236

This allele that has 561bp in-frame deletion in exon 14 of the *daf-2* gene that removes 187 amino acids from the TK domain of DAF-2 including the β-strands β1- β10 and the α-helices αC, αD and αE. This deletion removes both catalytic loop and most of the activation loop of the kinase domain (see Figure 5.5). No equivalent in frame deletions of the *h*IR have been reported. Potentially this allele has no kinase activity and may represent a molecular null. This allele and *m65* are investigated further in chapter 6.

Figure 5.5 Nonsense and deletion alleles



A schematic representation of the predicted translation products of one deletion and three nonsense alleles of *daf-2*. *m631* has an opal mutation at codon 281, which is in the L1 domain, the resulting protein product is predicted to be non-functional. *m646* has an ochre mutation at codon 1223, which is in the juxtamembrane region, this protein is also predicted to be non-functional. The *m65* allele has an amber mutation affecting codon 1449 in the TK domain; this protein is predicted to be kinase dead. The *tm1236* allele has a 187 amino acid deletion within the TK domain and is also predicted to result in a kinase dead protein.

5.4 Discussion

In this chapter I have sequenced 11 different alleles of *daf-2*, and located the lesion in 10 of them. All previously sequenced alleles have mutations either in the ligand-binding portion of the receptor (formed by the L1-CR-L2 domains) or the TK domain. However, in this project I have found mutations in four alleles that affect the Fn domains of the DAF-2 protein, as well as a *daf-2* allele with no mutation in its coding sequence. To gain further insight into the relationship between allele class and the molecular effect of each mutation we used homology modelling to predict the structural change at secondary, tertiary and quaternary levels. In addition, we performed a comparative analysis between our DAF-2 mutations and mutations in *h*IR and *h*IGF-1R, to gain further insight into the molecular effect of our lesions.

5.4.1 Distribution of mutations across DAF-2 receptor

An initial hypothesis, based on the sequence data available for several *daf-2* alleles (Kimura *et al.*, 1997; Scott *et al.*, 2002; Yu and Larsen, 2001), suggested that class 1 alleles had mutations affecting non-conserved residues in the L2 domain of the DAF-2 receptor, while class 2 alleles had mutations in the CR and TK domains of the receptor. A larger survey of the mutations in 24 alleles, including non-conditional mutants, suggests a more complex pattern for the location of class 1 and 2 alleles across the DAF-2 receptor. In general, class 1 alleles have mutations that affect the extracellular domain of the DAF-2 receptor, specifically the L2, Fn0 and Fn1 domains of the receptor. An exception is the atypical class 1 allele *e1369*, which has no mutation affecting the *daf-2* coding sequence and therefore a wild-type amino acid sequence. The majority of class 2 alleles have mutations affecting the CR and TK domains of the DAF-2 receptor, although the most severe class 2 allele *e979* has a mutation in the L1 domain and is the most N-terminal of the DAF-2 mutations. We had hoped to find that the mutation in some alleles affected the N- and C-terminal extensions, but none were found. This suggests that these regions may only have a minor role in DAF-2 function.

Table 5.4.1 Summary of the predicted molecular effects of mutations on DAF-2

daf-2 allele	Class	Phenotype	Molecular Effect Based on *h*IR and *h*IGF-1R mutants	Type of Defect
e1369	1	Severe Daf-c phenotype at 15°C	No mutation in coding sequence of *daf-2* gene	1
e979	2	Severe Daf-c phenotype at 15°C 100% Embryonic and L1 arrest at 25°C Eat and Unc	Involved in disulphide bond to C181 Reduced transport to cell surface Decreased insulin binding affinity	2, 3
m631	NC	Forms dauers at all temperatures	Nonsense allele that is predicted to form non-functional protein	1
m41	1	Strong Daf-c phenotype at 20°C Only allele to show maternal rescue of Daf-c	This residue is not present in non-nematode phyla	-
mg43	NC	Forms dauers at all temperatures	This allele has two missense mutations in DAF-2 CR domain First mutation involved in disulphide bond with C393 Both are predicted to impair proreceptor processing	2
m579	2	50% Daf-c phenotype at 22.5°C Eat and Unc	Reduced transport to cell surface Decreased insulin binding affinity Greater reduction in Shc phosphorylation than PI3 kinase phosphorylation Reduced receptor internalisation	2, 3, 4, 5
sa187	2	~100% Daf-c phenotype at 20°C Eat and Unc	Involved in disulphide bond to C483 Impaired transport to cell surface Decreased insulin binding affinity	2, 3
m596	2	Formerly class 2 on basis of Eat and Unc But suppressed by *daf-12(m20)*	Central part of a structural motif in L2 Impaired transport to cell surface No effect on insulin binding affinity No effect on autophosphorylation	2
e1368	1	Dauers recover at 25°C after 24 hours	Impaired transport to cell surface No effect on insulin binding affinity No effect on autophosphorylation	2
e1365/ sa193	1	Dauers recover at 25°C after 24 hours	Impaired transport to cell surface No effect on insulin binding affinity No effect on autophosphorylation	2
sa229	1	Dauers recover at 25°C after 24 hours	Accelerated degradation of receptor leading to a reduction of receptor levels at cell surface. No effect on insulin binding affinity No effect on autophosphorylation	5
mu150	1	Dauers recover at 25°C after 24 hours	Impaired transport to cell surface No effect on insulin binding affinity No effect on autophosphorylation	2
e1371	1	Dauers recover at 25°C after 24 hours	No known *h*IR and *h*IGF-1R mutants in equivalent region	2

m212	1	Severe Daf-c phenotype at 15°C	Involved with class I disulphide bond between α-subunits. Leads to reduction of holo-receptors at cell surface, as a small proportion of heterodimers fail to form tetramers. May reduce efficiency of transphosphorylation based on quaternary model	2, 4
m577	1	Only forms dauers at 25°C These do not recover	Involved in disulphide bond to C1055 Impaired transport to cell surface No effect on insulin binding affinity No effect on autophosphorylation	2
m646	NC	Forms dauers at all temperatures	Nonsense allele that is predicted to form non-functional protein	1
sa219	2	~100% Daf-c phenotype at 20°C Dauers recover after 8 days	Reduced insulin-stimulated autophosphorylation of kinase domain	4
e1391	2	Severe Daf-c phenotype at 20°C Eat and Unc	Reduced insulin-stimulated autophosphorylation of kinase domain	4
m65	NC	Forms dauers at all temperatures	Nonsense allele that is predicted to form non-functional protein	1
e1370	2	Canonical allele of *daf-2* Eat and Unc	Reduced insulin-stimulated autophosphorylation of kinase domain	4
m633	NC	Forms dauers at all temperatures	Reduced insulin-stimulated autophosphorylation of kinase domain	4
tm1236	NC	Forms dauers at all temperatures	Deletion of activation and catalytic loops of TK domain	4

Summary of the predicted molecular effects for each of the *daf-2* mutants. Alleles are listed in the order of their mutations across the DAF-2 amino acid sequence beginning with the N-terminus. They are classified into either class 1 or 2 if they are temperature-sensitive alleles or NC if they form dauers non-conditionally. The dauer phenotype and predicted molecular effects of each mutation are listed in columns 3 and 4. The final column uses the Taylor *et al* (1992) classification system to summarise the effects of each mutation (see Figure 1.4).

5.4.2 Possible molecular basis of the class 1 phenotype

The analysis described in this chapter has furnished much new information about the effects of *daf-2* mutations on the primary structure of the receptor and the predicted effects on secondary, tertiary and, in the case of *m212*, quaternary structure. Moreover, comparisons with mammalian work generate predictions about how these molecular defects affect receptor function. But how does this information help us understand the *daf-2* allele class difference at the molecular level? In the following discussion I survey the data generated, and attempt to draw from it broader conclusions about this question.

m596 is the only class 1 allele for which a mutation in the equivalent residue in *h*IR has been reported. This allele has a G546S substitution in the L2 domain of the DAF-2 receptor (Scott *et al.*, 2002), whilst the equivalent residue in *h*IR, G366, has an arginine substitution (Wertheimer *et al.*, 1994). The G366R mutation impairs the post-translation processing of the proreceptor, leading to a 90% reduction in levels of the mature receptor on the cell surface. This mutation did not affect the insulin-binding affinity of the mature receptor or its insulin-stimulated ability to autophosphorylate. If we assume that residues conserved between different members of the IR superfamily have similar structural roles within each receptor, then it seems likely that the G546S substitution in the *m596* mutant DAF-2 receptor also reduces receptor level on the cell surface by same mechanism. If, like the G366R *h*IR mutant, the G546S DAF-2 mutant receptor also retains wild-type ligand binding affinity and ability to autophosphorylate in response to ligand then potentially the phenotypes associated with the *m596* allele are due to a reduced flux of wild-type DAF-2 signalling output caused by the reduced receptor number on the cell surface. Thus a candidate molecular mechanism for the class 1 mutant defects is reduced flux of wild type DAF-2 signalling output caused by the reduced receptor number on the cell surface.

Do the predictions of the effects of the mutations in the other class 1 alleles support this hypothesis? We predict that the mutations in the *e1368*, *e1365/ sa193*, *mu150* and *e1371* alleles also will impair post-translational processing of the DAF-2 proreceptor in a similar way to the *m596* allele, and will also have reduced levels of mature DAF-2 on the cell surface. These five alleles are class 1A and form dauers that recover with 24 hours at 25.5°C (Garigan *et al.*, 2002; Gems *et al.*, 1998), unlike *m596*. This suggests that the postulated impairment of proreceptor processing is milder in these alleles than in *m596*.

By contrast, the above analysis implies that the mutation in the *sa229* allele leads to an accelerated rate of mature receptor internalisation and degradation leading to a reduced number of mature receptors on the cell surface, a similar effect is seen for the K460E and N462S mutants of *h*IR (Kadowaki *et al.*, 1990a; Kadowaki *et al.*, 1990c). The *sa229* allele is also class 1A (Gems *et al.*, 1998), and so the reduction in signalling flux may be similar to that seen in the other five class 1A alleles.

The mutation in the *m577* allele causes a C1045Y substitution; this cysteine is conserved across all members of the IR superfamily in our alignments and its equivalent in *h*IR C798 is known to be involved in a disulphide bond with C807 (Sparrow *et al.*,

1997), which also conserved across our alignment. Site-directed mutagenesis of the C798 residue and the equivalent C776 in *h*IGF-1R to serine results in impairment of proreceptor processing and lack of mature receptors on the cell surface (Maggi and Cordera, 2001).

The *m212* allele has severe Daf-c phenotype at 15°C and is one of two atypical class 1 alleles, the other being *e1369* (Gems *et al.*, 1998). Our analysis suggests that the mutation in *m212* will have two consequences. Firstly, it will lead to a severe impairment of the processing of the DAF-2 proreceptor reducing levels of mature receptor at the cell surface. Secondly, the loss of this disulphide bond will reduce the torsional resistance of the remaining bonds. This leads to a more random approach of the activation loops in the tyrosine kinase domains, reducing kinase activity. Thus, leading to a stronger reduction of signalling flux, which in turn leads to the stronger Daf-c phenotype.

The *e1369* allele is the most severe ts allele of *daf-2* in terms of its Daf-c phenotype, yet is class 1 (Gems *et al.*, 1998). This allele has no mutation in its coding sequence, so the receptor has a wild-type amino acid sequence. We believe that the mutation in *e1369* must affect a regulatory region of the *daf-2* gene, although our attempts to locate it so far have failed. This, we postulate, causes a severe decrease in the transcription of *daf-2* mRNA (see Chapter 6 for further investigation), which in turn leads to a highly reduced level of wild-type receptors on the cell surface and so a greatly reduced signalling flux.

In conclusion, our analysis of 9 distinct *daf-2* class 1 lesions shows that in every case there is evidence from the mammals of lowered receptor levels (see Table 5.4.1 for a summary). This is either a remarkable coincidence, or it implies that this is a common mechanism underlying the class 1 mutant phenotype.

5.4.3 Potential molecular basis of the class 2 phenotype

The *m579* and *e1391* alleles affect different parts of the DAF-2 receptor. The *m579* allele causes a R437C substitution in the CR domain (Scott *et al.*, 2002). This causes the loss of two hydrogen bonds with the E185 residue in the L1 domain; these bonds are crucial in keeping the L1 and CR domains in the correct orientation. The equivalent mutation R252C in *h*IR affects proreceptor processing leading to a reduced number

mature receptors on the cell surface (Hamer *et al.*, 2002). Unlike the G366R mutation in *h*IR, equivalent of the *m596* mutation, the R252C mutation also lowers the insulin binding affinity of the receptor and almost completely blocks insulin-stimulated internalisation of the receptor. However, the mutant receptor seems to undergo normal autophosphorylation and phosphorylates the IRS-1 and -2 proteins to wild-type level. But there is a marked reduction in the ability of this mutant to induce Shc and ERK1/2 phosphorylation and induce insulin-stimulated thymidine incorporation into DNA.

The *e1391* allele of *daf-2* has a P1434L substitution (Kimura *et al.*, 1997) in the TK domain. The equivalent mutation P1178L in *h*IR results in normal proreceptor processing and normal insulin binding affinity (Krook *et al.*, 1996). However the mutant receptor shows a marked decrease in insulin-stimulated autophosphorylation (Krook *et al.*, 1996; Krook *et al.*, 1997b; Moritz *et al.*, 1994), but in response to high insulin concentrations can phosphorylate IRS-1 to ~80% of wild-type receptor levels. However like the R252C mutation the P1178L mutant receptors also show a reduction in insulin-stimulated thymidine incorporation into DNA, they also have a severely reduced ability to phosphorylate the MAP kinase protein (homologous to ERK1/ ERK2) (Krook *et al.*, 1997b).

This implies that in both the *m579* and *e1391* alleles not only in there a reduction in the signalling flux through the DAF-2 receptor, as in the class 1 alleles, but this reduction seems to be asymmetric in that these alleles may have a greater impairment of the Shc/ MAP kinase pathway than the IST-1 (homologue of IRS-1/-2)/ PI3 Kinase pathway. The asymmetric nature of this reduction in signalling ability is achieved in different ways in the *h*IR equivalents of the two alleles. In the R252C mutant the reduced ability to phosphorylate Shc seems to be linked to the inability of the receptor to undergo insulin-stimulated internalisation, suggesting that receptor endocytosis is important for the Shc/ MAP kinase pathway (Hamer *et al.*, 2002). By contrast, in the P1178L mutant the reduced ability to phosphorylate MAP kinase seems to be linked to an alteration of the kinase activity of the TK domain itself.

Is this asymmetric reduction in signalling seen in the other class 2 alleles of *daf-2*? The *e979* allele has the most severe Daf-c phenotype of the class 2 alleles and has a 100% embryonic/ L1-stage arrest phenotype at 25.5°C (Gems *et al.*, 1998). The mutation in the *e979* allele is the most N-terminal mutation to affect the DAF-2 receptor and causes a C146Y mutation in the L1domain, which is predicted to affect a disulphide bond with C181 also in the L1 domain. Based on homology with several mutations in

the L1 domain of *h*IR, we expect the C146Y mutation in DAF-2 to not only impair processing of the proreceptor and thereby reduce the levels of the mature receptor at the cell surface, but to also affect the ligand binding affinity of the receptor and its ability to autophosphorylate in response to ligand-stimulation. We also expect that at 25.5°C the severity of the impairment of proreceptor processing increases, leading to a complete lack of receptors at the cell surface. This may explain the ts lethal phenotype of this allele.

These effects are consistent with the idea that all *daf-2* alleles have mutations that decrease the flux of DAF-2 signalling. However, we do not know how the predicted impaired autophosphorylation ability of the *e979* mutant receptor affects the phosphorylation of the Shc/ MAP kinase and IST-1/ PI3 kinase pathways. These latter remarks are also applicable to the *sa187* allele.

The *sa219* and *e1370* alleles both have mutations in the TK domain of the DAF-2 receptor, causing D1374N and P1465L substitution respectively (Kimura *et al.*, 1997). The P1465 residue is involved in a hydrogen bond with the R1430 residue in the P+1 loop of the TK domain based on the structures of the *h*IR and *h*IGF-1R (Favelyukis *et al.*, 2001; Hubbard *et al.*, 1994). We predict that the loss of this hydrogen bond in the P1465L mutation in the *e1370* allele will result in a similar but weaker impairment of kinase activity to that of the *e1391* allele. Thus, the *e1370* mutation is also predicted to cause an asymmetric reduction in signalling from DAF-2. As for the *sa219* allele there is no equivalent mutation in *h*IR, but we do predict it to also affect kinase activity.

5.4.4 Nonsense and deletion alleles

The *m646* allele has an ochre mutation at codon 1223 (K. Kimura and G. Ruvkun, personal communication). This allele is predicted to give rise to a DAF-2 receptor without a TK domain and so was expected to be a null allele. However, epistasis analysis of *m646* and two other non-conditional alleles of *daf-2* (*m65* and *m631*) with null alleles of *daf-16* and *daf-18* has revealed the following allelic series: *m65*< *m631*< *m646* in terms of DAF-2 activity (M. Nanji and D. Gems, unpublished results). We sequenced the lesions in the *m65* and *m631* to gain further insight into the nature of the DAF-2 defect in these alleles. The *m631* allele has an opal mutation at codon 281 in the L1 domain and *m65* has an amber mutation in codon 1449 of the TK domain.

The *m65* allele has the least DAF-2 activity, possibly none, of these three alleles. This result is surprising as the mutation in *m65* is the most C-terminal of the premature stop codons. Both *m631* and *m646* are predicted to produce non-functional protein, as these mutations should result in receptors lacking TK domains. In contrast, the *m65* mutation is located towards the C-terminus of the TK domain, and is predicted to give rise to a receptor with both the catalytic and activation loops of the receptor. The greater DAF-2 activity of *m631* and *m646* suggests that there is translational readthrough of the premature stop codons, and that these alleles give rise to some functional protein.

tm1236 is a deletion allele of *daf-2*. This allele was generated at our request by the Japanese *C. elegans* knockout consortium. The deletion removes 561bp of exon 14, leading to an 187aa deletion within the protein sequence. This deletion removes both the catalytic and activation loops of the receptor, and so is expected to be a molecular null. In the next chapter I further investigate both *m65* and *tm1236* in order to establish whether either allele or both represents a null allele of *daf-2*.

5.4.5 Concluding remarks

To gain further insight into phenotypic differences in *daf-2* mutants, we gathered data on the predicted mutational effects at the primary, secondary, tertiary and quaternary levels of protein structure, and the effect of equivalent mutations on *h*IR and *h*IGF-1R function. Overall, only the coupling of the primary sequence data and its comparison with equivalent or similar mutations in the other two receptors has furnished strong predictions about molecular basis of the *daf-2* allele class difference.

We find that class 1 alleles occur in the L2 and Fn domains, and class 2 alleles occur in the L1, CR and TK domains of the receptor. No mutations were found in the termini of the receptor. Comparison with *h*IR and *h*IGF-1R mutant data generates two predictions. Firstly, class 1 mutants have reduced levels of receptor at the cell surface, which retain wild-type function. Secondly, class 2 alleles may also have this defect but also have defects in signalling ability, which is often asymmetrical affecting Shc signalling more than PI3 kinase signalling. This suggests that Shc signalling could be involved with *daf-2*B function. Recently, it has been shown that the RAS pathway in *C. elegans* functions downstream of *daf-2* and may be involved with *daf-2*B function (Nanji *et al.*, 2005). This is consistent with reduced signalling via Shc, however, all this

is rather tentative. Predictions generated in this chapter will be explored further in the next chapter.

Chapter 6

Further analysis of *daf-2* alleles

6.1 Introduction

The sequencing of *daf-2* alleles described in Chapter 5 leads to a number of further hypotheses and predictions about the nature of the defects in individual *daf-2* mutants, and about the function of the wild type DAF-2 receptor. In this chapter I follow up on these predictions with a series of experimental tests. For example, I test whether the *daf-2(e1369)* allele causes reduced levels of transcription arising from a possible mutation in a regulatory element of the *daf-2* gene. I also attempted to use DAF-2 antibodies to confirm my predictions from the previous chapter about mutations expected to cause processing defects in the DAF-2 proreceptor. I investigate the correlation between DAF-2 signalling flux and class 1 and class 2 *daf-2* phenotypes through the use of RNAi. I have also used the *daf-28(sa191)* allele, which encodes a poisonous DAF-2 ligand, to probe the role of ligand-receptor interaction in determining class 1 and class 2 phenotypes. Finally I compare the properties of the *daf-2(m65)* nonsense allele to that of the *daf-2(tm1236)* deletion allele to determine if either represents a null allele.

6.1.1 Does *daf-2(e1369)* reduce gene transcription

In the previous chapter I showed that the atypical class 1 allele *e1369* did not contain a mutation in the CDS of the *daf-2* gene, and so encodes a wild-type DAF-2 receptor. In terms of its Daf-c phenotype the *e1369* allele is the most severe of all characterised ts alleles of *daf-2* (Gems *et al.*, 1998). Given that the *e1369* allele encodes a wild-type DAF-2 receptor the mutant phenotype of this allele could either be due to a mutation in a regulatory region of the *daf-2* gene affecting its transcription or a mutation in a different gene that is tightly linked to and phenocopies the *daf-2* gene. Sequencing of the most likely candidate regions for the mutation in *e1369*, the ~5.2kb upstream of the start codon and the entire first intron, did not identify any lesion. In order to check that the *e1369* allele has a mutation that affects a regulatory element of the *daf-2* gene and reduces its transcription I have used quantitative PCR (QPCR) to determine the expression level of the *daf-2* gene in an *e1369* mutant.

6.1.2 DAF-2 antibodies and processing defect detection

Mutations in *h*IR that cause defects in proreceptor processing can be detected through the use of antibodies that can detect the proreceptor and the cleaved α- and β-subunits of the mature receptor. In the previous chapter we predicted that several alleles of *daf-2* might have proreceptor processing defects based on homology modelling and comparison to mutations known to cause such defects in *h*IR. These defects may be detectable using a similar pulse-chase labelling assay with DAF-2 receptor antibodies. In addition to detecting processing defects, DAF-2 antibodies could also be used to address some more fundamental questions about the receptor. For example the DAF-2 receptor has large N- and C-terminal extensions, a feature shared with the *Cb*-DAF-2 and *Cr*-DAF-2 receptors (see Chapter 4, Table 4.3.9 and 4.3.10). The dInR receptor also has large extensions at both termini; the C-terminal extension is cleaved off in certain cells and retained in others (Fernandez *et al.*, 1995). It is unknown whether the N- and C-terminal extensions of DAF-2 are cleaved off, but it may be possible to detect this with an appropriate antibody.

6.1.3 DAF-2 signalling flux and the class 1-class 2 phenotype

In the previous chapter I presented the hypothesis that the class 1 alleles of *daf-2* have mutations that result in a reduction of wild-type signalling flux through the receptor, whereas class 2 alleles have mutations that result in an asymmetrical reduction of signalling flux through the receptor that affects the Shc/ MAP kinase pathway more than the PI3 kinase pathway. This asymmetry might be responsible for the different epistatic interactions of the two classes of *daf-2* alleles observed with genes such as *daf-12* (Gems *et al.*, 1998; Larsen *et al.*, 1995), *daf-9* (Gerisch and Antebi, 2004) and *rop-1* (Labbe *et al.*, 2000).

In order to further investigate the role of DAF-2 signalling flux in determining the class 1 and 2 phenotypes I used RNAi of *daf-2* to reduce signalling flux even more in both class 1 and 2 alleles. I also carried out *daf-2* RNAi in *daf-2* class 1 and 2 double mutants with *daf-12(m20)*. The *daf-12(m20)* allele is a hypomorphic nonsense mutation of the DAF-12 nuclear hormone receptor (Antebi *et al.*, 2000). This allele fully suppresses the Daf-c phenotype of class 1 *daf-2* alleles at 25.5°C and partially

suppresses the Age phenotype, whereas it leads to 100% embryonic and L1-stage arrest and enhancement of the Age phenotype in class 2 alleles (Gems *et al.*, 1998; Larsen *et al.*, 1995). The aim of these experiments was to see if further reduction of DAF-2 signalling flux could make class 1 alleles behave as class 2 alleles.

6.1.4 Ligand-receptor interaction

One possibility that might explain the differences between class 1 and class 2 alleles of *daf-2* with extracellular mutations is their effect on ligand binding. All class 1 alleles, except *e1369*, have mutations in the extracellular portion of the DAF-2 receptor. My earlier analysis indicates that these mutations do not affect ligand binding, suggesting that the receptor may remain sensitive to ligand stimulation. By contrast, class 2 alleles with mutations in the extracellular portion, such as *e979*, *m579* and *sa187*, are predicted to alter the ability of the receptor to bind ligand and in turn affect kinase activity.

The *daf-28(sa191)* allele encodes a mutant insulin-like peptide that appears to act as a poisonous ligand (i.e. antagonistic) of the DAF-2 receptor (Li *et al.*, 2003). If the above model is true, then one would expect that *daf-28(sa191)* would enhance the Daf-c phenotype of class 1 but not class 2 *daf-2* mutants. I this chapter I perform these tests and present the results.

6.1.5 Nonsense and deletion alleles of *daf-2* and the null phenotype

In the previous chapter I found that two non-conditional alleles of *daf-2*, *m65* and *m631*, contain nonsense mutations (Table 5.3.1). A third non-conditional allele *m646* had previously been identified as having a nonsense mutation (K. Kimura and G. Ruvkun, personal communication). Epistasis analysis of these three alleles with the null alleles *daf-16(mgDf50)* (Ogg *et al.*, 1997) and *daf-18(nr2037)* (Mihaylova *et al.*, 1999) places them in the following order in terms of increasing phenotypic severity *m646* < *m631* < *m65* (M. Nanji and D. Gems, unpublished data) (Table 6.1.1). The *daf-18(nr2037)* allele also fully suppresses the Age phenotype of *m646* and *m631*, but not *m65*, although Age in the latter is suppressed by *daf-16(mgDf50)* (M. Nanji and D. Gems, unpublished data). This implies that *m646* and *m631* retain some DAF-2 activity, most likely due to

translational read-through of the premature stop codons. The early larval arrest phenotype of *daf-2(m65)* is only partially suppressed by either *daf-16(mgDf50)* or *daf-18(nr2037)* at 15°C and the *daf-16(mgDf50)*; *daf-2(m65)* double mutants shows 100% embryonic lethality at 25°C. A similar phenotype has been previously reported with *daf-16(m26)*; *daf-2(m65)* double mutants (Larsen *et al.*, 1995). *m65* may be a null allele of *daf-2*.

The Japanese *C. elegans* gene knockout consortium isolated the *tm1236* allele of *daf-2*, after our lab requested it. This allele contains a 561bp in-frame deletion within exon 14 of the *daf-2* gene, which translates into a 187 amino acid deletion of a part of the TK domain (see Chapter 5 for more details). This allele is predicted to be kinase dead and so also may represent a null allele of *daf-2*. I compared the phenotypes of the *m65* and *tm1236* alleles both individually and in combination with *daf-16(mgDf50)* to test which is more severe (and therefore more likely to be null), and to gain further insight into *daf-2* function.

Table 6.1.1 Incomplete suppression of *daf-2* nonsense alleles by *daf-16(0)* and *daf-18(0)*

Genotype	% L4 and Adults	% L1 and L2	% Dauers	% Dead eggs	N
25°C					
N2	99	0	0	1	295
daf-18(nr2037)	100	0	0	0	275
daf-16(mgDf50)	100	0	0	0	325
daf-2(m65); daf-18	24	6	0	70	171
daf-2(m631); daf-18	67	26	0	7	85
daf-2(m646); daf-18	92	5	0	3	283
daf-16; daf-2(m65)	0	0	0	100	298
daf-16; daf-2(m646)	100	0	0	0	172
15°C					
N2	99	0	0	1	230
daf-18(nr2037)	100	0	0	0	189
daf-16(mgDf50)	100	0	0	0	257
daf-2(m65); daf-18	*65*	6	5	24	109
daf-2(m631); daf-18	*78*	0	22[a]	0	172
daf-2(m646); daf-18	89	0	11[a]	0	221
daf-16; daf-2(m65)	*43*	18	0	39	292
daf-16; daf-2(m646)	100	0	0	0	210

Summary of epistasis data for three nonsense alleles of *daf-2* with null alleles of *daf-16* and *daf-18*. Only 65% and 24% of *daf-2(m65); daf-18(nr2037)* double mutants reach adulthood at 15°C and 25°C, respectively. A higher proportion of animals reach adulthood in both *daf-2(m631); daf-18(nr2037)* and *daf-2(m646); daf-18(nr2037)* double mutants at these temperatures. This suggests that these nonsense alleles vary in their levels of IIS pathway activity, which in turn suggests that they have some residual DAF-2 activity due to translation read through. The data implies that *m646* has the weakest impairment of DAF-2 activity followed by *m631* and then *m65*. Data from *daf-16(mgDf50); daf-2(m65)* and *daf-16(mgDf50); daf-2(m646)* double mutants also supports this view. [a] Partial dauers, recovered within 24 hr. **Data from M. Nanji and D. Gems (unpublished results)**

6.2 Materials and Methods

6.2.1 Strains

Strains used for QPCR of *e1369*

The following strains were used in the QPCR experiments; GA202 *daf-16(mgDf50) I*; *daf-2(m212) III*, GA309 *daf-16(mgDf50) I*; *daf-2(e1369) III*, GR1307 *daf-16(mgDf50) I* (Ogg *et al.*, 1997), DR608 *daf-2(m212) III* and DR1573 *daf-2(e1369)* III. Both of the *daf-16*; *daf-2* double mutants were created by crossing males of the parental *daf-2* strain with *daf-16(mgDf50)* hermaphrodites at 15°C. The hermaphrodite progeny of these crosses were then shifted to fresh plates at 25°C as gravid adults and allowed to lay eggs for 24 hours before being discarded. After three days at 25°C the plates were screened dauers and approximately 50 dauers were picked to fresh plates for each cross and shifted to 15°C to allow recovery. All dauers were *daf-2* homozygotes, some of which would be *daf-16*(+) homozygotes and others *daf-16*(+)/ (-) heterozygotes. The recovered dauers were plated singly and shifted to 25°C once they had begun egg laying. Again the worms allowed to lay eggs for 24 hours before being discarded. Once again the plates were left at 25°C before being screened. Plates that contained 100% dauer progeny were discarded and adults were picked from plates with a mixture of dauers and adults. These adults were *daf-16*; *daf-2* double mutants.

Strains used for antibody and RNAi experiments

The following strains were used for the antibody work N2, GA158 *daf-16(mgDf50) I*; *daf-2(m65) III* and SU159 *ajm-1(ok160) X*; *jcEx44* (Koppen *et al.*, 2001). The strains used in the RNAi experiments were DR20 *daf-12(m20) X*, DR1296 *daf-2(e1370) III*; *daf-12(m20) X*, DR1547 *daf-2(m41) III*; *daf-12(m20) X*, DR1563 *daf-2(e1370) III* and DR1564 *daf-2(m41) III*.

Strains used for construction of *daf-2*; *daf-28* double mutants

The strains used in the *daf-2*; *daf-28* double mutant experiments were CB1068 *unc-79(e1068) III*, BE148 *rol-9(sc148) V*, DR1566 *daf-2(m579) III*, DR1567 *daf-2(m577) III*, GA304 *unc-79(e1068) III*; *daf-28(sa191) V*, GA9 *daf-2(m577) III*; *rol-9(sc148) V*, GA199 *daf-2(m579) III*; *rol-9(sc148) V*, GA305 *daf-2(m577) III*; *daf-28(sa191) V*; GA306 *daf-2(m579) III*; *daf-28(sa191) V* and JT191 *daf-28(sa191) V* (Li *et al.*, 2003).

In order to construct the *daf-2*; *daf-28* doubles I first constructed double mutants for both *daf-2* alleles with *rol-9(sc148)*, which was used as marker for *daf-28* as it is within 2cM of this gene. The *rol-9(sc148)* allele has a recessive roller (Rol) phenotype. I crossed males of either parental *daf-2* strain to *rol-9(sc148)* hermaphrodites at 15°C. The hermaphrodite progeny of these crosses were then shifted to fresh plates at 25°C as gravid adults. They were allowed to lay eggs for 24 hours being discarded. After three days at 25°C the plates were screened for Rol dauers, approximately 20 were picked for each *daf-2* allele and shifted to fresh plates at 15°C to allow recovery. These adults were *daf-2*; *rol-9* double mutants.

The second step in the construction of the *daf-2*; *daf-28* doubles was the construction of the *unc-79(e1038)*; *daf-28(sa191)* double mutant. The *unc-79* gene was used a marker for the *daf-2* gene as it is within 7cM of this gene, this allele has a recessive uncoordinated movement (Unc) phenotype. *daf-28(sa191)* males were crossed with *unc-79(e1038)* hermaphrodites at 15°C and the hermaphrodite progeny of this cross were shifted to 25°C as gravid adults and allowed to lay eggs for 24 hours. The *daf-28(sa191)* allele is semi-dominant and approximately 10% of *sa191*/ + animals form dauers at 25°C (Malone and Thomas, 1994). For this reason I picked ~20 Unc adults after three days at 25°C, which were then transferred singly to fresh plates and left at 25°C. These adults were allowed to lay eggs for 24 hours before they were discarded. After three days the plates were scored for dauers, plates with no dauers were discarded. Approximately 30 dauers were picked to fresh plates and shifted to 15°C to aid recovery. Once these animals had become gravid adults they were plated singly and shifted to 25°C and allowed to lay eggs for 24 hours before being discarded. *daf-28(sa191)* homozygotes form 99% dauers at 25°C (Malone and Thomas, 1994), so plates with less than ~100% dauers after three days were discarded. Several animals

from plates with ~100% dauers were picked and shifted to 15°C to aid recovery. These animals were *unc-79(e1038)*; *daf-28(sa191)* double mutants.

The final step in the construction of the *daf-2*; *daf-28* doubles involved crossing *unc-79(e1038)*; *daf-28(sa191)* males with *daf-2*; *rol-9(sc148)* hermaphrodites at 15°C. The hermaphrodite progeny were then shifted to fresh plates at 25°C and allowed to lay eggs for 24 hours before being discarded. The plates were then screened for dauers after three days, ~50 non-Rol non-Unc dauers were picked to fresh plates and left at 25°C for a further day. The *daf-28(sa191)* allele forms dauers that recover at 25°C (Malone and Thomas, 1994), whereas *daf-2(m577)* and *daf-2(m579)* dauers do not recover at 25°C (Gems *et al.*, 1998). The plates were then scored again and animals that had remained dauers were picked to fresh plates and transferred to 15°C to allow recovery. At this point these dauers could either be *daf-2* homozygotes; *daf-28/ rol-9* heterozygotes, assuming no recombination had occurred between *daf-28* and *rol-9*, or *daf-2*; *daf-28* homozygotes. The recovered animals were shifted to fresh plates singly at 25°C as gravid adults and allowed to lay eggs for 24 hours before being discarded. Plates were scored three days later and plates that segregated Rol and non-Rol dauers were discarded. Plates segregating 100% non-Rol dauers were transferred back to 15°C to allow recovery. These animal were assumed to be *daf-2*; *daf-28* homozygotes, which was confirmed by sequencing their genomic DNA for the presence of both lesions (data not shown).

Strains used for nonsense and deletion allele experiments

The strains used in the *daf-2* nonsense and deletion allele work were GA8 *mDf11 III/ qC1* [*dpy-19(e1259), glp-1(q339)*] *III*, GA158 *daf-16(mgDf50) I*; *daf-2(m65) III*, GA176 *daf-16(mgDf50) I*; *mDf11 III/ qC1* [*dpy-19(e1259) III, glp-1(q339)*] *III*, GA301 *daf-2(tm1236) III/ +*, GA302 *daf-2(tm1236) III / qC1* [*dpy-19(e1259) III* and GA310 *daf-16(mgDf50) I*; *daf-2(tm1236) III.*

The *daf-2(tm1236)* allele was originally shipped as a heterozygote over wild-type by the Japanese *C. elegans* knockout consortium, as their notes said the homozygotes were lethal and/ or sterile. This strain had no official designation so I labelled it GA301. I found that the *daf-2(tm1236)/ +* heterozygotes segregated 25% dauer progeny at 15°C and 20°C (data not shown). Also the deletion allele and the wild type allele could be

detected using single worm PCR with the *daf-2* exon 14-specific primers F16F and Ex14F2a (Table 5.2.1) (data not shown).

I then balanced the *tm1236* allele over the *qC1* [*dpy-19(e1259) III, glp-1(q339)*] inversion by crossing *mDf11 III/ qC1* males with putative *daf-2(tm1236)/ +* heterozygotes at 20°C. The *mDf11* deficiency (Gems *et al.*, 1998) is a large deletion within *C. elegans* chromosome III that deletes the *daf-2* gene as well as several others, although the exact break points have not been characterised. When homozygous the *mDf11* deficiency is lethal (dead eggs). The *qC1* inversion has two recessive markers that when homozygous results in a swollen body 'dumpy' (Dpy) phenotype and at 25°C has a germline-proliferative mutation that makes the worms sterile (Glp). Putative *daf-2(tm1236)/ +* were plated singly and allowed to lay eggs for 24 hours. The parental worms were then checked for the presence of the *tm1236* deletion allele using single worm PCR. Plates that had progeny laid by worms that were not heterozygotes for the *tm1236* deletion allele were discarded. The progeny of the cross were raised to adulthood and then ~20 were plated singly at 25°C and allowed to lay eggs for 24 hours. The plates were kept at 25°C for three days before being scored, plates segregating dauers, Dpy Glp and wild-type worms were kept and the rest discarded. Several wild-type worms from these plates were plated singly at 25°C and allowed to lay eggs for 24 hours before being discarded. The progeny were then scored to ensure they segregated as 25% dauers, 25% Dpy Glp and 50% wild-type animals. The wild-type animals were *daf-2(tm1236)/ qC1*. These animals were then backcrossed twice to the *mDf11 III/ qC1* strain.

To construct the *daf-16(mgDf50)*; *daf-2(tm1236)* double mutants I crossed *daf-16(mgDf50)*; *mDf11/ qC1* males with *daf-2(tm1236)/ qC1* hermaphrodites at 20°C, and then shifted the mated hermaphrodites to 25°C. The worms were allowed to lay eggs for 24 hours before being discarded. The only fertile progeny of this cross would be *daf-16(mgDf50)/ +*; *daf-2(tm1236)/ qC1* and *daf-16(mgDf50)/ +*; *mDf11/ qC1* animals. Approximately 20 fertile hermaphrodites from this cross were plated singly at 25°C and allowed to lay eggs for 24 hours before being discarded. Plates that did not segregate any dauers after three days at 25°C were discarded, as their parents would have been *daf-16(mgDf50)/ +*; *mDf11/ qC1* animals. Plates with dauers were retained and ~20 non-dauer non- Dpy Glp animals were picked and plated singly at 25°C, these animals were allowed to lay eggs for 24 hours before being discarded. Plates were scored three days later. Plates that segregated dauers and Dpy Glp animals were discarded, as were plates

that segregated Dpy Glp animals but no dauers. Plates that segregated 100% fertile adults were retained and several animals were transferred to 20°C for 24-hour egg lays. This worms were then assayed for the presence of both the *daf-2(tm1236)* deletion and the *daf-16(mgDf50)* deletion. All worms were found to be homozygous for both deletions and so were *daf-16(mgDf50)*; *daf-2(tm1236)* double mutants (data not shown).

6.2.2 *e1369* Quantitative PCR

Analysis of the *daf-2* transcript level in the *daf-2(e1369)* mutant was carried out using RNA from *daf-16(mgDf50)*; *daf-2(e1369)* animals. These double mutants were used because they can be grown more easily than the *daf-2(e1369)* single mutant, which has a severe Daf-c phenotype even at 15°C (Gems *et al.*, 1998). Two different control genotypes were included in the QPCR experiments the first was *daf-16(mgDf50)*, which has a wild type *daf-2* gene. The second was the *daf-16(mgDf50)*; *daf-2(m212)* double mutant, *m212* is the second most severe allele after *e1369*, both are class 1, also the mutation in *m212* is known (Table 5.3.1).

Three RNA samples were prepared for each genotype using the methods described in Chapter 3. Protocols for the first strand cDNA synthesis and QPCR can be found in section 3.2.2 and 3.2.4 of the main materials and methods chapter. The *ama-1* gene was used as the endogenous control in all experiments. This gene encodes the large subunit of the RNA polymerase II protein (Bird and Riddle, 1989) and has been used as a control gene in quantitative transcript analysis previously as its post-embryonic mRNA abundance is not subject to strong fluctuation (Johnstone and Barry, 1996; Larminie and Johnstone, 1996).

Primers for *ama-1* transcripts were designed using the Primer3 WWW interface (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) (Rozen and Skaletsky, 2000). These primers span exon-exon boundaries in order to prevent the amplification of any contaminating genomic DNA within the RNA samples. The *ama-1*Fwd primer spans the exon 1-exon 2 boundary and the *ama-1*Rev primer spans the exon 2-exon 3 boundary. The primers for the *daf-2* transcripts were F2F and F1R (Table 5.2.1), which anneal to exon 3 and exon 4 of the *daf-2* gene, respectively, which are separated by a 257bp intron (Table 5.3.4). The primer sequences are listed in Table 6.2.1.

Table 6.2.1 Primer pairs used in QPCR

Primer name	5'- Sequence - 3'
*ama-1*Fwd	CCCGGAGGAGATTAAACGCATG
*ama-1*Rev	CATCTTCCACGACGATCTATGA
daf-2 F2F	TGAGACGGCTTCGACTTGTCA
daf-2 F1R	TGGCGGGTTTTCCTCATAGCA

Each plate had the general layout seen in Table 6.2.2. All of the QPCR experiments were performed using the relative standard curve method (Applied Biosystems user bulletin 2, http://docs.appliedbiosystems.com/pebiodocs/04303859.pdf).

6.2.3 DAF-2 antibodies and Western blots

Two different polyclonal antibodies were used in the western blot analysis of the DAF-2 receptor. The first was a commercially available antibody, cN-17 (Santa Cruz Biotechnology Inc.), raised against a peptide mapping to the N-terminus of the DAF-2 receptor. The second DAF-2 antibody, anti-DAF-2 (Genosphere Biotechnologies), was generated against a peptide mapping to the C-terminal extension of the receptor. The peptide was designed by me and has the sequence 'CLTNRGGSNERGAGF' equivalent to residues 1679-1693 of the DAF-2 receptor. The monoclonal MH27 antibody (Francis and Waterston, 1991) was used as control antibody and detects the AJM-1 protein from *C. elegans* (Koppen *et al.*, 2001). The MH27 antibody was obtained from the Developmental Studies Hybridoma Bank at the University of Iowa.

The secondary antibodies used in the western analysis were as follows polyclonal anti-goat IgG, monoclonal anti-goat IgG, polyclonal anti-rabbit IgG, monoclonal anti-rabbit IgG, polyclonal anti-mouse IgG and Fc-specific anti-mouse IgG (all from Sigma). All antibodies were peroxidase conjugates for use in chemiluminescence assays. Protocols for protein gels and western analysis can be found in Chapter 3.

Table 6.2.2 General QPCR plate layout

	1	2	3	4	5	6	7	8	9	10	11	12
A	NTC *ama-1*	NTC *ama-1*	NTC *ama-1*	1xN2 *ama-1*	1xN2 *ama-1*	1xN2 *ama-1*	1/2 N2 *ama-1*	1/2 N2 *ama-1*	1/2 N2 *ama-1*	1/4 N2 *ama-1*	1/4 N2 *ama-1*	1/4 N2 *ama-1*
B	1/16 N2 *ama-1*	1/16 N2 *ama-1*	1/16 N2 *ama-1*	1/64 N2 *ama-1*	1/64 N2 *ama-1*	1/64 N2 *ama-1*	1/128 N2 *ama-1*	1/128 N2 *ama-1*	1/128 N2 *ama-1*	*mgDf50*/1 *ama-1*	*mgDf50*/1 *ama-1*	*mgDf50*/1 *ama-1*
C	*mgDf50*/2 *ama-1*	*mgDf50*/2 *ama-1*	*mgDf50*/2 *ama-1*	*mgDf50*/3 *ama-1*	*mgDf50*/3 *ama-1*	*mgDf50*/3 *ama-1*	*mgDf50*; *m212*/1 *ama-1*	*mgDf50*; *m212*/1 *ama-1*	*mgDf50*; *m212*/1 *ama-1*	*mgDf50*; *m212*/2 *ama-1*	*mgDf50*; *m212*/2 *ama-1*	*mgDf50*; *m212*/2 *ama-1*
D	*mgDf50*; *m212*/3 *ama-1*	*mgDf50*; *m212*/3 *ama-1*	*mgDf50*; *m212*/3 *ama-1*	*mgDf50*; *e1369*/1 *ama-1*	*mgDf50*; *e1369*/1 *ama-1*	*mgDf50*; *e1369*/1 *ama-1*	*mgDf50*; *e1369*/2 *ama-1*	*mgDf50*; *e1369*/2 *ama-1*	*mgDf50*; *e1369*/2 *ama-1*	*mgDf50*; *e1369*/3 *ama-1*	*mgDf50*; *e1369*/3 *ama-1*	*mgDf50*; *e1369*/3 *ama-1*
E	NTC *daf-2*	NTC *daf-2*	NTC *daf-2*	1xN2 *daf-2*	1xN2 *daf-2*	1xN2 *daf-2*	1/2 N2 *daf-2*	1/2 N2 *daf-2*	1/2 N2 *daf-2*	1/4 N2 *daf-2*	1/4 N2 *daf-2*	1/4 N2 *daf-2*
F	1/16 N2 *daf-2*	1/16 N2 *daf-2*	1/16 N2 *daf-2*	1/64 N2 *daf-2*	1/64 N2 *daf-2*	1/64 N2 *daf-2*	1/128 N2 *daf-2*	1/128 N2 *daf-2*	1/128 N2 *daf-2*	*mgDf50*/1 *daf-2*	*mgDf50*/1 *daf-2*	*mgDf50*/1 *daf-2*
G	*mgDf50*/2 *daf-2*	*mgDf50*/2 *daf-2*	*mgDf50*/2 *daf-2*	*mgDf50*/3 *daf-2*	*mgDf50*/3 *daf-2*	*mgDf50*/3 *daf-2*	*mgDf50*; *m212*/1 *daf-2*	*mgDf50*; *m212*/1 *daf-2*	*mgDf50*; *m212*/1 *daf-2*	*mgDf50*; *m212*/2 *daf-2*	*mgDf50*; *m212*/2 *daf-2*	*mgDf50*; *m212*/2 *daf-2*
H	*mgDf50*; *m212*/3 *daf-2*	*mgDf50*; *m212*/3 *daf-2*	*mgDf50*; *m212*/3 *daf-2*	*mgDf50*; *e1369*/1 *daf-2*	*mgDf50*; *e1369*/1 *daf-2*	*mgDf50*; *e1369*/1 *daf-2*	*mgDf50*; *e1369*/2 *daf-2*	*mgDf50*; *e1369*/2 *daf-25*	*mgDf50*; *e1369*/2 *daf-2*	*mgDf50*; *e1369*/3 *daf-2*	*mgDf50*; *e1369*/3 *daf-2*	*mgDf50*; *e1369*/3 *daf-2*

NTC= No template control

6.2.4 *daf-2* RNAi

For RNAi of the *daf-2* gene, HT115(DE3) cells transformed with a construct containing 587 bp of exon 9 of the *daf-2* gene ligated into the L4440 vector were used (kindly provided by Adam Antebi). HT115(DE3) cells transformed with the L4440 empty vector were used as the control treatment. A high penetrance lactose-based induction method was used for all assays, described in Chapter 3.2.11. All worm strains were grown for at least three generations on RNAi bacteria at 15°C before the Daf-c assays were performed.

6.2.5 Daf-c assays

Daf-c assays for the strains undergoing *daf-2* RNAi and for the nonsense and deletion allele strains were carried out at two different temperatures 20°C and 25°C. Assays for the *daf-2*; *daf-28* double mutants were carried out at 15°C, 20°C, 22.5°C and 25°C. All Daf-c assay were carried out as follows: 10 gravid hermaphrodites were transferred from 15°C on to fresh plates and allowed to lay eggs for six hours at the relevant temperature. In most cases there were two plates per strain at each temperature. The adults were then either discarded, or transferred to fresh plates at the same temperature that they had been laying eggs at. The progeny were then scored for dauer formation at the following time points: 48 hours for animals at 25°C, 72 hours for animals at 22.5°C, 80 hours for animals at 20°C and 96 hours for animals at 15°C.

6.3 Results

6.3.1 *daf-2(e1369)* transcript levels

The *e1369* allele of *daf-2* has no mutation affecting the coding sequence of the gene or the UTRs of the mRNA transcript (see Chapter 5), despite having the most severe Daf-c phenotype of any ts allele. I first confirmed that this allele is a *daf-2* mutant by performing complementation tests with two different mutants for which the sequence change is known: *m41* (Yu and Larsen, 2001) and *m577* (this project, see Chapter 5). The *e1369* allele fails to complement either of these alleles (data not shown). This excludes the possibility that *e1369* lies in a closely linked gene with a *daf-2*-like phenotype.

One possible explanation for the phenotype of this allele is the presence of a mutation in a regulatory element of the *daf-2* gene. Sequencing of ~5.2kb of genomic DNA upstream of the start codon and the entire ~5.5kb first intron of the *daf-2* gene failed to locate the mutation. Given the size of the *daf-2* gene, ~33kb, and the size of the 5' intergenic space, ~14kb (see Chapter 4), direct sequencing of the remaining 37kb of genomic DNA as a means to identify the *e1369* mutation was considered too laborious and expensive for my thesis project. If the *e1369* allele does have a mutation in a regulatory element it may be possible to detect its effect on the levels of the *daf-2* transcript using QPCR.

Table 6.3.1 Relative *daf-2(e1369)* transcript levels

Trial	Relative abundance in *daf-16(mgDf50)* ± sd	Relative abundance in *daf-16(mgDf50)*; *daf-2(m212)* ± sd	Relative abundance in *daf-16(mgDf50)*; *daf-2(e1369)* ± sd
1	1 ± 0.77	21.54 ± 0.98	4.74 ± 1.51
2	1 ± 1.26	1.99 ± 0.38	0.76 ± 0.51
3	1 ± 0.85	1.18 ± 0.67	1.79 ± 0.62
4	1 ± 0.63	1.77 ± 0.76	4.52 ± 0.83
5	1 ± 0.74	1.35 ± 0.54	0.50 ± 0.77
6	1 ± 1.29	1.06 ± 1.32	1.32 ± 0.83
7	1 ± 0.83	0.34 ± 0.62	0.23 ± 0.41
8	1 ± 0.59	1.13 ± 0.49	1.03 ± 0.72
9	1 ± 0.54	1.26 ± 0.35	0.86 ± 0.47
10	1 ± 0.47	0.59 ± 0.42	0.29 ± 0.41

Relative levels of *daf-2* transcript in *daf-16(mgDf50)*, *daf-16(mgDf50)*; *daf-2(m212)* and *daf-16(mgDf50)*; *daf-2(e1369)* animals across 10 independent trials. There is no clear evidence for a decrease in *daf-2* transcript in *e1369* animals.

The QPCR experiments were carried with *daf-16(mgDf50)*; *daf-2(e1369)* double mutants as these worms are Daf-d and can easily be grown in large, stage-synchronous cultures, unlike *daf-2(e1369)* worms which will arrest as 100% dauers at high population density. *daf-16(0)*; *daf-2(+)* worms were used as the wild-type control. *daf-16(0)*; *daf-2(m212)* double mutants were included as second control for the general effects of daf-2 mutations per se on *daf-2* mRNA levels. The *m212* allele is the second most severe allele of *daf-2* after *e1369* and its mutation has been identified (see Chapter 5) and is not predicted to alter *daf-2* transcript levels.

In total 10 different QPCR experiments have been performed to measure the relative transcript levels of the *daf-2* gene in *daf-16(0)*, *daf-16(0)*; *daf-2(e1369)* and *daf-16(0)*; *daf-2(m212)* worms (Table 6.3.1). The levels of *daf-2* transcript in *daf-16(mgDf50)*; *daf-2(m212)* were similar to the levels in *daf-16(mgDf50)* in five trials. In two trials it was almost double the *daf-16(mgDf50)* levels, although the standard deviations overlap so these results are not significant. In the first trial the levels of *daf-2* transcript in *daf-16(mgDf50)*; *daf-2(m212)* was 22-fold greater than the *daf-16(mgDf50)*. There were two trials in which levels were lower than *daf-16(mgDf50)*, but again the standard deviations overlap so these results are not significant.

6.3.2 Using DAF-2 antibodies to detect processing defects

In the previous chapter a combination of homology modelling and comparisons with both naturally occurring and induced mutations of the *h*IR and *h*IGF-1R implied that several mutants of DAF-2 might cause defects in the processing of the proreceptor. It has been possible to detect such defects in *h*IR by performing pulse-chase labelling experiments with cultured cells and observing the various stages in the maturation of the receptor with the use of polyclonal antibodies, first demonstrated with the H209R mutant of *h*IR by Kadowaki *et al* (1991).

In order to try and test my predictions about mutations that cause similar processing defects in the DAF-2 receptor, I tested two different DAF-2 antisera using western blots. The first was a commercially available goat polyclonal antibody, cN-17, from Santa Cruz Biotechnology Inc. This antibody was raised against a peptide that maps to the N-terminus of the DAF-2 protein, although the exact sequence of the peptide is not available from the company.

The second antibody, which I will refer to as anti-DAF-2, was a rabbit polyclonal antibody that was raised by Genosphere Biotechnologies against a peptide that I designed. This peptide had the sequence 'CLTNRGGSNERGAGF' and maps to residues 1679-1693 of the DAF-2 receptor. This sequence lies within the C-terminal extension of the receptor. I chose a to raise an antibody to peptide mapping to the C-terminus of DAF-2 for two reasons. The first was the fact that an antibody against the N-terminus was already available. The second reason was to see whether the C-terminal extension of DAF-2 is cleaved off, like the C-terminal extension of *D. melanogaster* dInR (Fernandez *et al.*, 1995).

A requirement for the selection of the 15aa peptide sequence was the presence of a cysteine residue as either the N- or C-terminal amino acid to allow conjugation to the KLH carrier protein. There is only one cysteine residue present in the C-terminal extension of DAF-2 at position 1679. Analysis of the antigenicity of the 14 amino acids on either side of C1679 using the Protean™ module of the Lasergene suite suggested either portion of sequence would be suitable. I decided to use the sequence C-terminal of C1679 as this region is better conserved in *Cb*-DAF-2 than the sequence N-terminal of C1679 (Figure 6.1). The hope was that the antiserum might also recognize *Cb*-DAF-2. This region is also better conserved in *Cr*-DAF-2.

Figure 6.1 Alignment of peptide sequence used to raise anti-DAF-2

```
DAF-2       TGASTSSYTGGGPYCLTNRGGSNERGAGF
Cb-DAF-2    TGASNCSYTG-GPYCLTNRGGSNER-AGF
Cr-DAF-2    TGASTASYAG-GPY-LANRGGSNER-AGF
            ****..**:* *** *:******** ***
```

Peptide chosen for antibody. Selection highlighted in green is the region chosen for raising antiserum. The cysteine residue for conjugation to KLH carrier protein shown in red.

The first attempt at western blotting with the two different DAF-2 antibodies was carried out using 1:1000 dilutions of each antibody and 1:5000 dilutions of the relevant polyclonal secondary antibody. I used the MH27 antibody as a positive control; this is a monoclonal antibody that recognizes the AJM-1 protein in *C. elegans* (Koppen *et al.*, 2001). Protein extracts were made from N2 and *daf-16(mgDf50)*; *daf-2(m65)* worms. The *m65* allele of *daf-2* is nonsense mutant that has a stop codon at position 1449 of the DAF-2 amino acid sequence. This allele is a candidate null and its mutation is predicted to remove the C-terminal extension of the DAF-2 protein. The anti-DAF-2 antibody was raised against a peptide that maps to the C-terminal extension. Therefore a western blot with this antibody should show a band in the N2 protein sample that is absent in the *daf-16(mgDf50)*; *daf-2(m65)* protein samples. The cN-17 antibody was raised against a peptide that maps to the N-terminus of the DAF-2 receptor. This antibody should detect bands in both protein samples, however the band in the *daf-16(mgDf50)*; *daf-2(m65)* protein samples should have a lower molecular weight as it lacks 398aa from the C-terminus of the receptor.

My first blot using the conditions above detected multiple bands that were present in both the N2 and the *daf-16(mgDf50)*; *daf-2(m65)* protein samples. In addition these bands were identical in blots for each antibody (Figure 6.2). I suspected that these bands were not being recognized by my primary antibodies, but were instead being bound by the secondary antibodies. I therefore repeated the blot, but omitted the primary antibodies and used only the secondary antibodies at the same dilution of 1:5000. The blot contained the same bands that had been present in my first blot confirming that the bands seen were the result of binding by the secondary antibodies (data not shown). One possible interpretation of this finding is that the secondary antibodies were raised in animals with nematode infections.

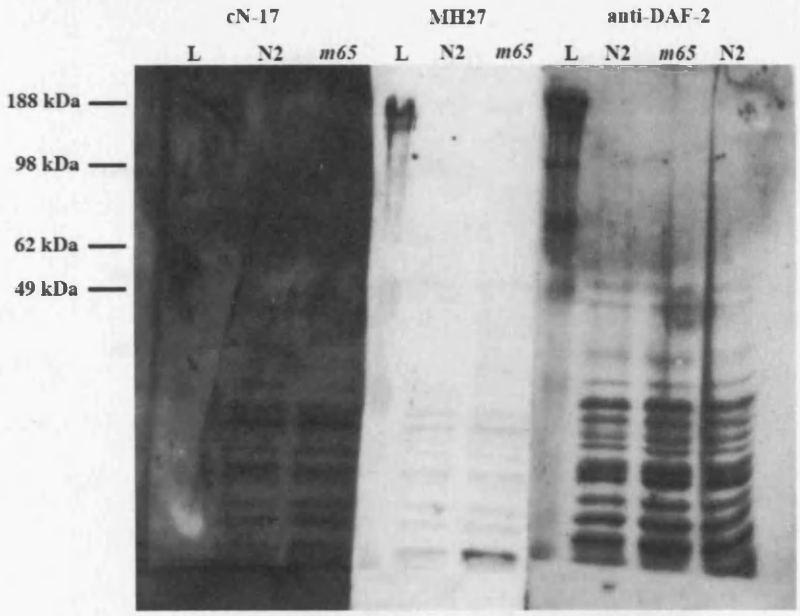
The secondary antibodies used above were polyclonal peroxidase-linked antibodies from Sigma. In an attempt to prevent the detection of these spurious bands I switched to using monoclonal anti-Goat IgG and anti-Rabbit IgG antibodies as the secondary antibodies for the DAF-2 antibodies (Sigma). I used an anti-Mouse IgG Fc-specific antibody (Sigma) as the secondary antibody to bind the MH27 antibody. I carried out a blot in which only these secondary antibodies were used, again at a dilution of 1:5000. These antibodies did not detect any bands in either the N2 or the *daf-16(mgDf50)*; *daf-2(m65)* protein samples (Figure 6.3). However, for unknown reasons both monoclonal anti-Goat IgG and anti-Rabbit IgG antibodies bound to the molecular weight ladder.

Using the new secondary antibodies I performed another blot. The dilution of the primary antibodies was 1:1000 and the secondary antibody dilutions were kept at 1:5000. The two different DAF-2 antibodies were incubated with protein samples from N2 and *daf-16(mgDf50)*; *daf-2(m65)* worms. The MH27 antibody was only incubated with N2 protein extract. The MH27 antibody detected multiple bands (Figure 6.4). The cN-17 antibody also detected several bands. However, these bands were the same size in both the N2 and *daf-16(mgDf50)*; *daf-2(m65)* protein samples. These bands are therefore unlikely to correspond to the DAF-2 protein. The anti-DAF-2 antibody did not bind to anything in either protein sample (Figure 6.4). I repeated the western blots several more times using the same conditions as above and also with primary antibody dilutions of 1:500. These blots displayed the same bands as seen in Figure 6.4 (data not shown). Suggesting that neither of the DAF-2 antisera contain significant levels of anti-DAF-2 antibodies, at least not antibodies that bound to DAF-2 under conditions in which the MH27 positive control antibody bound its complementary antigen.

The MH27 monoclonal antibody recognizes the AJM-1 protein (Koppen *et al.*, 2001). The *ajm-1* gene encodes five different splice isoforms of the AJM-1 protein. Previous analysis of the MH27 antibody suggests that its antigen is greater than 150KDa in weight (Francis and Waterston, 1991). Because my western blots produced multiple bands with the MH27 antibody when using protein samples from N2 worms I decided to try using protein samples from worms with a deletion allele of *ajm-1* to ensure that MH27 was actually working as a positive control. The SU159 strain is homozygous for the *ajm-1(ok160)* deletion allele. This mutation is embryonic lethal, so the SU159 strain carries a *ajm-1*::*gfp* fusion construct as an extra-chromosomal array that rescues the lethal phenotype of this mutant (Koppen *et al.*, 2001). This construct

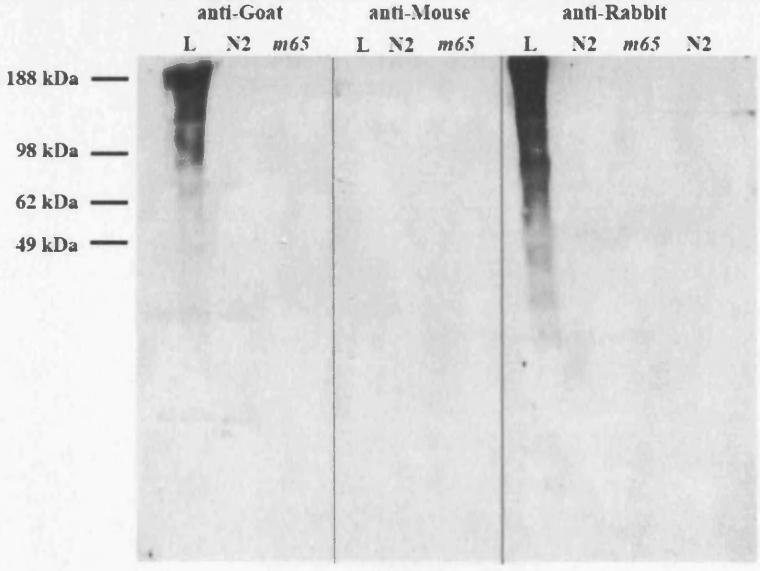
was made using genomic DNA that includes exons 2-8 of the *ajm-1* gene, which are common to all splice isoforms of this gene. Western blots using this strain are missing two bands compared to protein samples from N2 worms, which are likely to correspond to the two largest isoforms of AJM-1, but have a very bright additional band that corresponds to the AJM-1::GFP fusion protein (Figure 6.5). This confirms that the conditions used for western blotting were capable of generating signal with anti-nematode antibodies. This in turn provides support for the view that neither DAF-2 antiserum contained sufficiently high levels of DAF-2 antibodies to be useful.

Figure 6.2 1st attempt at western blotting with DAF-2 antibodies



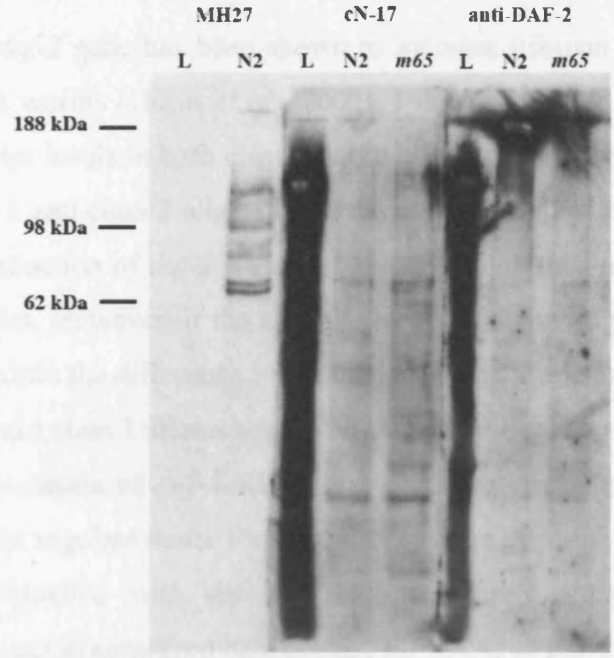
First attempt at western blot. L is the molecular weight ladder. N2 is wild-type protein sample. *m65* is *daf-16(0)*; *daf-2(m65)* protein sample. cN-17 is the Santa Cruz DAF-2 antibody, MH27 is the control antibody and anti-DAF-2 is the Genosphere DAF-2 antibody. Note binding to similar bands in all three blots when using polyclonal secondary antibodies.

Figure 6.3 Western with new secondary antibodies only



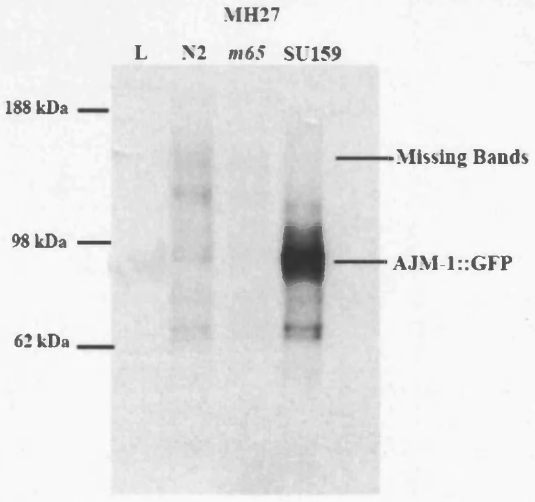
Abbreviations as above. Note that the anti-Goat and anti-Rabbit monoclonal antibodies bind the molecular weight ladder.

Figure 6.4 Western blot using primary antibodies with new secondary antibodies



Western blot using monoclonal secondary antibodies. L is the molecular weight ladder. N2 is wild-type protein sample. *m65* is *daf-16(0)*; *daf-2(m65)* protein sample. Note MH27 binds multiple bands, the bands bound by cN-17 are the same in both protein samples, and the anti-DAF-2 antibody did bind anything.

Figure 6.5 Western blot of N2 and SU159 protein sample with MH27



Abbreviations as above. Note that two of the bands bound by MH27 in the N2 protein sample are missing in the SU159 protein sample, which is a *ajm-1* deletion allele, but carries a rescuing transgenic *ajm-1*::*gfp* fusion construct . These are likely to be the largest isoforms of AJM-1.

6.3.4 Effect of *daf-2* RNAi in class 1 and class 2 alleles

RNAi of the *daf-2* gene has been shown to increase lifespan by reducing its transcript levels in adult worms (Dillin *et al.*, 2002). I decided to use RNAi of *daf-2* to further reduce transcript levels in both class 1 and 2 alleles. If my prediction that the difference between class 1 and class 2 alleles is the asymmetric reduction of Shc signalling is true then further reduction of *daf-2* transcript levels should have no effect on the phenotype of class 1 alleles. However, if the absolute levels of signalling flux through the receptor is what determines the difference between class 1 and 2 alleles then potentially RNAi of *daf-2* could make class 1 alleles behave as class 2.

The two classes of *daf-2* allele have differing epistatic interactions with several other genes that regulate dauer formation. Given the opposite effects of the two classes when in combination with *daf-12(m20)* I wondered if *daf-2*; *daf-12(m20)* double mutants might act as sensitized background for RNAi of *daf-2*. Potentially class 1 *daf-2*; *daf-12* double mutants, which are normally Daf-d at 25°C, might develop the embryonic and larval arrest phenotype of class 2 *daf-2*; *daf-12* double mutants.

Table 6.3.2 class 1 and 2 RNAi at 20°C Trial 1

Genotype	RNAi Treatment	% Dauers	% Dead Eggs	% L1 & L2	% Young adult-arrested animals	% Late larval & Adults	N
daf-12(m20)	L4440	0	0	0	0	100	399
daf-12(m20)	*daf-2*	0	0	0	0	100	329
daf-2(m41)	L4440	100	0	0	0	0	298
daf-2(m41)	*daf-2*	100	0	0	0	0	133
daf-2(m41); daf-12(m20)	L4440	0	0	0	0	100	312
daf-2(m41); daf-12(m20)	*daf-2*	0	0	0	0	100	188
daf-2(e1370)	L4440	6.2	0	0	0	93.8	227
daf-2(e1370)	*daf-2*	100	0	0	0	0	35
daf-2(e1370); daf-12(m20)	L4440	0	0	0	0	100	188
daf-2(e1370); daf-12(m20)	*daf-2*	0	0	2.0	98.0	0	49[a]

Effects of *daf-2* RNAi in class 1 (*m41*) and class 2 (*e1370*) alleles of *daf-2* as single mutants and as *daf-2*; *daf-12* double mutants at 20°C. *daf-2* RNAi has no effect on *m41* in either the single or double mutant, whereas *e1370* forms 100% dauers and *daf-2(e1370); daf-12(m20)* double mutants arrest as sterile young adults. This effect is replicated in the second trial below. [a] In both trials the total number of worms was coincidently the same

Table 6.3.3 class 1 and 2 RNAi at 20°C Trial 2

Genotype	RNAi Treatment	% Dauers	% Dead Eggs	% L1 & L2	% Young adult-arrested animals	% Late larval & Adults	N
daf-12(m20)	L4440	0	0	3.1	0	96.9	293
daf-12(m20)	*daf-2*	0	0	0.6	0	99.4	308
daf-2(m41)	L4440	97.4	0	2.6	0	0	117
daf-2(m41)	*daf-2*	100	0	0	0	0	43
daf-2(m41); daf-12(m20)	L4440	0	0	0	0	100	96
daf-2(m41); daf-12(m20)	*daf-2*	0	0	0	0	100	88
daf-2(e1370)	L4440	0	0	0	0	100	93
daf-2(e1370)	*daf-2*	0	0	100*	0	0	43
daf-2(e1370); daf-12(m20)	L4440	0	0	0.9	0	99.1	114
daf-2(e1370); daf-12(m20)	*daf-2*	0	0	2.0	98.0	0	49[a]

[a] In both trials the total number of worms was coincidently the same * Arrested as L2d

Table 6.3.4 class 1 and 2 RNAi at 25°C Trial 1

Genotype	RNAi Treatment	% Dauers	% Dead Eggs	% L1 & L2	% Young adult-arrested animals	% Late larval & Adults	N
daf-12(m20)	L4440	0	0	1.1	0	98.9	360
daf-12(m20)	*daf-2*	0	0	1.0	0	99.0	397
daf-2(m41)	L4440	99.2	0	0.8	0	0	385
daf-2(m41)	*daf-2*	86.1	3.0	10.9	0	0	101
daf-2(m41); daf-12(m20)	L4440	0	0.5	0	0	99.5	398
daf-2(m41); daf-12(m20)	*daf-2*	0	0	8.6	91.4	0	151
daf-2(e1370)	L4440	92.5	0	7.5	0	0	228
daf-2(e1370)	*daf-2*	98.1	0	1.9	0	0	52
daf-2(e1370); daf-12(m20)	L4440	0	0	100	0	0	140
daf-2(e1370); daf-12(m20)	*daf-2*	0	0	100	0	0	12

Effects of *daf-2* RNAi in class 1 (*m41*) and class 2 (*e1370*) alleles of *daf-2* as single mutants and as *daf-2*; *daf-12* double mutants at 25°C. *daf-2* RNAi has no effect on *daf-2(m41)*. However, it causes the majority of *daf-2(m41)*; *daf-12(m20)* animals to arrest as sterile young adults, similar to the arrest seen in *daf-2(e1370)*; *daf-12(m20)* animals at 20°C. This effect is replicated in the second trial below. *daf-2* RNAi has no effect on *daf-2(e1370)* or *daf-2(e1370)*; *daf-12(m20)* animals at this temperature.

Table 6.3.5 class 1 and 2 RNAi at 25°C Trial 2

Genotype	RNAi Treatment	% Dauers	% Dead Eggs	% L1 & L2	% Young adult-arrested animals	% Late larval & Adults	N
daf-12(m20)	L4440	0	0	2.1	0	97.9	340
daf-12(m20)	*daf-2*	0	0	0	0	100	341
daf-2(m41)	L4440	96.2	0	3.8	0	0	132
daf-2(m41)	*daf-2*	100	0	0	0	0	55
daf-2(m41); daf-12(m20)	L4440	0	1.42	13.5	0	85.1	141
daf-2(m41); daf-12(m20)	*daf-2*	0	0	11.8	88.2	0	76
daf-2(e1370)	L4440	97.3	0	2.8	0	0	109
daf-2(e1370)	*daf-2*	82.4	0	17.7	0	0	34
daf-2(e1370); daf-12(m20)	L4440	0	0	100	0	0	55
daf-2(e1370); daf-12(m20)	*daf-2*	0	0	100	0	0	18

RNAi of the *daf-2* gene was carried out in the following backgrounds *daf-12(m20)*, *daf-2(m41)*, *daf-2 (e1370)*, *daf-2(m41)*; *daf-12(m20)* and *daf-2(e1370)*; *daf-12(m20)*. Two different temperatures were used (20°C and 25°C) and the lactose-based induction method (see Chapter 3). The trials at 20°C were included as controls to check for the efficacy of *daf-2* RNAi, as the only observable phenotype of this construct on wild-type worms is lifespan extension, which was not measured in these experiments. In preliminary trials at this temperature *daf-2(e1370)* did not form dauers on L4440 control cells but formed 100% dauers on *daf-2* RNAi (data not shown).

At 20°C RNAi of *daf-2* in *daf-2(m41)* (class 1) was uninformative, since both L4440 control and *daf-2* RNAi worms formed 100% dauers (Table 6.3.2 and 6.3.3). On standard NGM plates seeded with OP50 this allele forms ~14% dauers (Gems *et al.*, 1998). Possibly the lactose plates lead to increased dauer formation in an *m41* background, although this did not enhance dauer formation in *e1370* (Table 6.3.2 and 6.3.3). RNAi of *daf-2* in an *e1370* background led to 100% dauer formation in the first trial and 100% L2d formation in the second trial, confirming that there is knockdown of *daf-2* levels with the use of RNAi in these trials.

RNAi of *daf-2* in the *daf-2*; *daf-12* strains had no effect on the *m41* double mutant: as the animals developed directly into adults. However *daf-2* RNAi has an interesting effect on the *e1370* double mutant. These worms arrest as sterile young adults with no gonadal development, and shifting these worms to 15°C and onto L4440 plates does not rescue this effect (data not shown). This phenotype is seen with *daf-2(e1370)*; *daf-12(m20)* mutants grown at 22.5°C on standard NGM plates with OP50 (Gems *et al.*, 1998; Larsen *et al.*, 1995). The fact that this arrest is not seen in the L4440 control worms suggests that it is due to the RNAi of *daf-2*. There also appears to be a reduction in fecundity of *e1370* doubles but not *m41*.

RNAi of *daf-2* at 25°C had no effect on *daf-12(m20)* worms in either trial. Again there was little effect on *daf-2(m41)*; although in the first trial there were ~10% L1/ L2 arrest this was not replicated in the second trial. RNAi of *daf-2* in *e1370* also had no detectable effect, although the worms on *daf-2* RNAi laid fewer eggs than control worms. RNAi of the *e1370* double mutants also seemed to reduce fecundity of the worms, compared to controls, but did not have any other effects. Critically RNAi of *daf-2* in the *m41* double mutants in both trial resulted in retarded larval development and ultimate arrest as sterile young adults with no gonadal development, similar to that seen in the *e1370* double mutants at 20°C. Thus, RNAi of *daf-2* changes the behaviour of the

class 1 *m41* mutant into that of a class 2 mutant. This implies that overall levels of signalling flux through the DAF-2 receptor is a determinant of whether an allele behaves as class 1 or class 2.

6.3.5 Analysis of *daf-2*; *daf-28* double mutants

The *C. elegans* genome has 39 genes that encode insulin-like peptides, *ins-1* to *-37* (Duret *et al.*, 1998; Gregoire *et al.*, 1998; Kawano *et al.*, 2000; Kawano *et al.*, 2003; Pierce *et al.*, 2001), *ins-38* (M. Costa, personal communication) and *daf-28* (Li *et al.*, 2003). The role of these peptides remains unclear given that DAF-2 is the only member of the IR superfamily encoded in the worm genome, but it is postulated that some of these peptides are DAF-2 agonists and others antagonists.

e979, *m579* and *sa187* are all class 2 *daf-2* alleles with mutations in the extracellular portion of the receptor. These alleles are predicted to have altered ligand-binding affinity, based on similar mutations in *h*IR (chapter 5), which in turn is expected to alter their kinase activity. Class 1 alleles all have extracellular mutations, except *e1369*, but unlike the extracellular class 2 alleles these mutations are not predicted to affect ligand binding and may remain responsive to ligand.

The *daf-28(sa191)* allele is the only mutant of an insulin-like peptide in *C. elegans* that has a phenotype (Li *et al.*, 2003). This allele is semi-dominant resulting in ~10% dauer formation in heterozygous animals (Malone and Thomas, 1994). When homozygous this allele forms dauers at all temperatures, with increasing severity as temperature increases, which are able to recover (Malone and Thomas, 1994). The *sa191* allele has a missense mutation causing a R37C substitution that affects the proteolytic processing site of the DAF-28 peptide (Li *et al.*, 2003). The mutant DAF-28 peptide is believed to act as poisonous ligand that interferes with signalling role of the other INS peptides.

I have constructed *daf-2*; *daf-28(sa191)* double mutants with both class 1 (*m577*) and class 2 (*m579*) alleles of *daf-2*. If the above is correct, *sa191* should enhance the mutant phenotype of class 1, but not class 2 mutants. I have performed Daf-c assays with the following genotypes; N2, *daf-28(sa191)*, *daf-2(m577)*, *daf-2(m579)*, *daf-2(m577)*; *daf-28(sa191)* and *daf-2(m579)*; *daf-28(sa191)*. Because the *sa191* allele forms a proportion of dauers at all temperatures I performed the Daf-c assays at 15°C,

20°C, 22.5°C and 25°C. It should be noted that the first trials at each temperature were set up on a day when temperatures in our lab exceeded 30°C due to an air conditioning failure; this may be account for the increased embryonic and early larval arrest seen in the first trial compared to the second trial at each temperature.

Table 6.3.6 *daf-2; daf-28* Daf-c at 15°C Trial 1

Genotype	% Dauers	% Dead Eggs	% L1 & L2	% Late larval & Adults	N
N2	0	0	5.0	95.0	120
daf-28(sa191)	0	0	10.5	89.5	76
daf-2(m577)	0	0	6.2	93.8	96
daf-2(m579)	0	0	5.0	95.0	81
daf-2(m577); daf-28(sa191)	39.2	0	16.2	44.6	74
daf-2(m579); daf-28(sa191)	0	0	7.0	93.0	43

Dauer formation in class 1(*m577*) and class 2 (*m579*) *daf-2*; *daf-28* double mutants at 15°C. The *m577* allele does not form dauers at this temperature. The *daf-2(m577)*; *daf-28(sa191)* double mutant forms ~40% dauers in both trials at 15°C, which means that *daf-28(sa191)* enhances the Daf-c phenotype of this allele. The *daf-2(m579)*; *daf-28(sa191)* double mutant behaves similarly to the *daf-28(sa191)* mutant in both trials, suggesting that the *m579* allele is unresponsive to the *daf-28(sa191)* poison ligand.

Table 6.3.7 *daf-2; daf-28* Daf-c at 15°C Trial 2

Genotype	% Dauers	% Dead Eggs	% L1 & L2	% Late larval & Adults	N
N2	0	0	1.7	98.3	351
daf-28(sa191)	8.5	0	0	91.5	271
daf-2(m577)	0	0	0.7	99.3	283
daf-2(m579)	0	0	0	100	260
daf-2(m577); daf-28(sa191)	37.9	0	0	62.1	132
daf-2(m579); daf-28(sa191)	13.0	0	0	87.0	170

At 15°C *daf-28(sa191)* has been reported to form 8.1% dauers (Malone and Thomas, 1994). In these Daf-c experiments at 15°C the *sa191* allele formed no dauers in the first trial, although there was ~10% L1 and L2 arrest (Table 6.3.6), and 8.5% dauers

in the second trial (Table 6.3.7). The *m577* and *m579* alleles form no dauers at this temperature. Addition of the *sa191* allele increased dauer formation in *m577* but not *m579* mutants in both trials. This supports the view that in *m577* but not *m579* mutants, the DAF-2 receptor remains responsive to ligand.

At 20°C the *sa191* allele formed between 30-40% dauers in the two trials. The *m577* allele forms no dauers at this temperature. As at 15°C, the Daf-c phenotype of the *daf-2(m577); daf-28(sa191)* was greater than the single mutants. The *m579* allele was previously been shown to form between 0% (Gems *et al.*, 1998) and 4.2% (Labbe *et al.*, 2000) dauers at 20°C. Here, at 20°C the *m579* allele formed no dauers in the first trial, and ~39% dauers in the second trial (Table 6.3.9). At 20°C *sa191* allele also enhanced the Daf-c phenotype of *m579* in the double mutants. However, the level of dauer formation in the *daf-2(m579); daf-28(sa191)* strain was lower than in *daf-2(m577); daf-28(sa191)*. This is despite the fact that *m577* is a much weaker allele in terms of Daf-c than *m579* (Gems *et al.*, 1998) (Table 6.3.10 and 6.3.11). This implies that the DAF-2 receptor in the *m579* mutation can still respond to the *sa191* poison ligand.

At 22.5°C the *sa191* allele again formed between 30-40% dauers in the two trials (Table 6.3.10 and 6.3.11). The *m577* allele formed 0-0.5% dauers, consistent with previous observations in Gems *et al* (1998). Again the *sa191* allele enhances the Daf-c phenotype of the *m577* allele in the double mutant. The embryonic arrest seen in the double mutants, respectively, in the first trial is probably due to temperature problems described earlier, since the second trial does not have any embryonic arrest at this temperature. The *m579* allele formed ~90% and ~98% dauers in the two trials at 22.5°C. This is higher than previous observations of 50% (Gems *et al.*, 1998) and 60.9% (Labbe *et al.*, 2000) at this temperature. At this temperature an interaction between *m579* and *sa191* was undetectable due to the high level of Daf-c resulting from *m579* alone.

The *sa191* allele was previously found to form 99% dauers at 25°C (Malone and Thomas, 1994). In my two trials the *sa191* allele formed ~71% and ~97% dauers (Table 6.3.12 and 6.3.13). *m577* allele formed ~80% and ~90% dauers in the two trials, with ~10% L1/ L2 arrest in both trials. The *daf-2(m577); daf-28(sa191)* double mutants formed ~31% and ~71% dauers, with ~62% and ~13% embryonic arrest, respectively. The high degree of embryonic arrest in the first trial was most likely due to the temperature problems described before. The *m577* allele on it own did not display embryonic arrest at 25°C in the second trial, but the *m577; sa191* double did show a significant level of such arrest. Possibly this means that the *daf-2(m577); daf-28(sa191)*

double mutant is behaving as class 2 mutant of *daf-2*, which often show some embryonic arrest at 25°C (Gems *et al.*, 1998). The *daf-2(m579)*; *daf-28(sa191)* double mutants also show enhanced embryonic arrest compared to the *m579* and *sa191* alleles on their own. Casual observations suggest that both double mutants also have drastically reduced fecundity compared to the single mutants.

To further test the possibility that *daf-2(m577)*; *daf-28(sa191)* behaves as a class 2 mutant I tested the *daf-2*; *daf-28* double mutants for the class 2 25°C Unc phenotype. Both *daf-2(m579)* and *daf-2(m579)*; *daf-28(sa191)* alleles developed the Unc phenotype by the third day at this temperature (data not shown). In contrast the *daf-2(m577)*, *daf-28(sa191)* and *daf-2(m577)*; *daf-28(sa191)* mutants had not become Unc by the sixth day of monitoring. Taken together results from these *daf-2*; *daf-28* studies provide support for the model derived from the sequencing and modelling studies of Chapter 5. It appears that class 1 mutants are more responsive to ligand to INS ligand than class 2 mutations affecting the extracellular domain.

Table 6.3.8 *daf-2; daf-28* Daf-c 20°C Trial 1

Genotype	% Dauers	% Dead Eggs	% L1 & L2	% Late larval & Adults	N
N2	0	0	1.9	98.1	208
daf-28(sa191)	27.4	0	3.1	69.5	164
daf-2(m577)	0	0	3.3	96.7	153
daf-2(m579)	0	0	8.1	91.9	197
daf-2(m577); daf-28(sa191)	87.1	0	6.0	6.9	101
daf-2(m579); daf-28(sa191)	75.0	0	25.0	0	36

Table 6.3.9 *daf-2; daf-28* Daf-c 20°C Trial 2

Genotype	% Dauers	% Dead Eggs	% L1 & L2	% Late larval & Adults	N
N2	0	0	0	100	470
daf-28(sa191)	40.9	0	0	59.1	342
daf-2(m577)	0	0	0	100	380
daf-2(m579)	38.7	0	0	61.3	225
daf-2(m577); daf-28(sa191)	97.4	0	0	2.6	151
daf-2(m579); daf-28(sa191)	67.4	0	0.7	31.9	138

Dauer formation in class 1(*m577*) and class 2 (*m579*) *daf-2*; *daf-28* double mutants at 20°C. The *m577* allele does not form dauers at this temperature. The *daf-2(m577)*; *daf-28(sa191)* double mutant forms between ~90-100% dauers in both trials at 20°C, the enhancement of the Daf-c phenotype is much stronger than at 15°C (Table 6.3.6 and 6.3.7). The *daf-2(m579)*; *daf-28(sa191)* double mutant also shows an enhancement of the Daf-c phenotype at this temperature.

Table 6.3.10 *daf-2; daf-28* Daf-c 22.5°C Trial 1

Genotype	% Dauers	% Dead Eggs	% L1 & L2	% Late larval & Adults	N
N2	0	0	0	100	201
daf-28(sa191)	39.9	0	8.1	52.0	148
daf-2(m577)	0.5	0	4.50	95.0	219
daf-2(m579)	89.6	7.2	3.2	0	221
daf-2(m577); daf-28(sa191)	61.2	35.5	3.3	0	121
daf-2(m579); daf-28(sa191)	78.4	17.7	3.9	0	51

Table 6.3.11 *daf-2; daf-28* Daf-c 22.5°C Trial 2

Genotype	% Dauers	% Dead Eggs	% L1 & L2	% Late larval & Adults	N
N2	0	0	0	100	242
daf-28(sa191)	31.9	0	0	68.1	157
daf-2(m577)	0	0	0	100	164
daf-2(m579)	97.8	0	2.2	0	136
daf-2(m577); daf-28(sa191)	97.9	0	2.1	0	95
daf-2(m579); daf-28(sa191)	96.2	0	3.8	0	105

Dauer formation in class 1(*m577*) and class 2 (*m579*) *daf-2*; *daf-28* double mutants at 22.5°C. The *m577* allele does not usually form dauers at this temperature. The *daf-2(m577)*; *daf-28(sa191)* double mutant has a strong embryonic lethal phenotype in the first trial (35.5%) that is not displayed in the second trial. However, the Daf-c phenotype of *m577* is still enhanced by *daf-28(sa191)*. At this temperature both *daf-2(m579)* and *daf-2(m579)*; *daf-28(sa191)* form ~100% dauers, which obscures any possible interaction between the two genes.

Table 6.3.12 *daf-2; daf-28* Daf-c 25°C Trial 1

Genotype	% Dauers	% Dead Eggs	% L1 & L2	% Late larval & Adults	N
N2	0	0	3.4	96.6	146
daf-28(sa191)	71.4	24.4	4.2	0	119
daf-2(m577)	80.9	10.9	8.2	0	183
daf-2(m579)	86.8	3.3	9.9	0	151
daf-2(m577); daf-28(sa191)	30.8	61.7	7.5	0	107
daf-2(m579); daf-28(sa191)	16.0	60.0	24.0	0	25

Table 6.3.13 *daf-2; daf-28* Daf-c 25°C Trial 2

Genotype	% Dauers	% Dead Eggs	% L1 & L2	% Late larval & Adults	N
N2	0	0	2.5	97.5	243
daf-28(sa191)	97.0	0	0.6	2.4	169
daf-2(m577)	90.4	0	9.6	0	156
daf-2(m579)	84.6	1.9	13.5	0	156
daf-2(m577); daf-28(sa191)	71.3	13.4	15.3	0	157
daf-2(m579); daf-28(sa191)	29.2	70.8	0	0	48

Dauer formation in class 1(*m577*) and class 2 (*m579*) *daf-2*; *daf-28* double mutants at 25°C. The *daf-2(m577)*; *daf-28(sa191)* double mutant has a strong embryonic lethal phenotype in the first trial, 61.7%, but only 13.4% in the second trial. This trait is seen in class 2 alleles of *daf-2*, potentially the *daf-2(m579)*; *daf-28(sa191)* double mutant is phenocopying a class 2 allele. The *daf-2(m579)*; *daf-28(sa191)* double mutant also has a strong embryonic lethal phenotype in both trials.

6.3.6 Analysis of *daf-2(m65)* and *daf-2(tm1236)*

Currently there are no fully confirmed null alleles of *daf-2*, although Gems *et al* (1998) provide evidence that *m65* is a strong candidate. We have sequenced the *m65* coding sequence and identified its lesion (Chapter 5, Table 5.3.1). *m65* contains a nonsense mutation with an amber stop codon replacing the W1449 codon. Two other nonsense alleles of *daf-2* have also been identified: these are *m631* (this thesis) and *m646* (K. Kimura and G. Ruvkun, personal communication). The stop codons in both *m631* and *m646* are more N-terminal than that of *m65* (Table 5.3.1). The *m631* allele has opal stop codon at position 281 of the DAF-2 amino acid sequence and so the predicted protein produced by this allele will be truncated at the L1 domain and lack all activity. The *m646* allele has an ochre codon at position 1223 of the DAF-2 amino acid sequence and its predicted protein would lack the entire TK domain and C-terminal extension of the receptor. The *m65* mutation affects the C-terminal half of the TK domain, specifically the αF helix that follows on from the catalytic and activation loops of the TK domain. So the predicted protein produced by this allele would contain two of the most important parts of the TK domain.

Based on the position of the nonsense codons one might infer that *m631* and *m646* represent better candidates for a *daf-2* null allele. However, epistasis analysis of these three nonsense mutants with null alleles of *daf-16* and *daf-18* does not support this prediction (M. Nanji and D. Gems, unpublished data) (Table 6.1.1). The *m646* allele is fully suppressed by both *daf-16(mgDf50)* and *daf-18(nr2037)* at 15°C and 25°C, whereas the *m631* allele displays ~30% L1 and L2 arrest at 25°C in a *daf-18(nr2037)* background. The *m65* allele produces 100% and 70% dead eggs at 25°C in *daf-16(0)* and *daf-18(0)* backgrounds, respectively. This suggests the following allelic series *m65*< *m631*< *m646* based on DAF-2 activity. The greater suppression of *m646* and *m631* by *daf-18(0)* compared to *m65* implies that these alleles must have some DAF-2 activity, even though their predicted protein sequences lack the TK domain, and so are not null. Most likely there is some translational readthrough of the *m631* and *m646* transcripts producing functional protein. It remains unclear whether there is readthrough of *m65* transcript and therefore it remains the best candidate for a *daf-2* null allele.

Recently the Japanese *C. elegans* knockout consortium has isolated a deletion allele of *daf-2*. The *tm1236* allele has an in-frame 561bp deletion of exon 14, which leads to a 187 amino acid deletion within the TK domain of the DAF-2 receptor (Chapter 5). This deletion removes both the catalytic and activation loops of the TK domain, which are both present in the *m65* predicted protein, but retains the C-terminal extension of the DAF-2 receptor. Therefore the *tm1236* allele is predicted to be kinase dead and is a good candidate for a null allele of *daf-2* on molecular grounds.

I decided to assay the *m65* and *tm1236* alleles to see if it was possible to distinguish between the two phenotypically and establish if either or both are null alleles of *daf-2*. I performed Daf-c assays on both alleles either balanced over the *qC1* inversion, as both alleles are non-conditional, or as *daf-16(mgDf50)*; *daf-2* double mutants at 25°C. This temperature was used because of the 100% embryonic lethal phenotype observed in *daf-16(mgDf50)*; *daf-2(m65)* double mutants by Manoj Nanji and David Gems (personal communication), a similar phenotype was found with *daf-16(m26)*; *daf-2(m65)* double mutants (Larsen *et al.*, 1995).

In the first of two trials at 25°C I found that the *m65* and *tm1236* alleles, when over the *qC1* balancer, segregated 23.0% and 26.0% dauers, respectively (Table 6.3.14). Neither of these alleles had any embryonic lethality at this temperature when over the balancer chromosome. When these alleles were placed in the *daf-16(0)* background at this temperature I found that *tm1236* allele was almost fully suppressed, although it did have ~7% L1 and L2 arrest. Unexpectedly I found that the *daf-16(0)*; *daf-2(m65)* double mutant produced only 31.3% dead eggs, with 44.8% of animals arresting as L1s and L2s. Interestingly, I found that 23.96% of animals made it through to adulthood.

In the second trial at 25°C I found that only 12.9% and 15.8% of progeny arrested as dauers from the *m65/ qC1* and *tm1236/ qC1* worms, respectively (Table 6.3.15). However, this was accompanied by a rise in the number of worms arrested as L1s and L2s. In this trial I found that the majority of *daf-16(0)*; *daf-2(tm1236)* double mutants reached adulthood, while 12.8% arrested as L1s and L2s. The *daf-16(0)*; *daf-2(m65)* double mutants produced 42.2% dead eggs, 22.3% L1s and L2s and 35.5% reached adulthood. Possibly, the discrepancies between this trial and the first were due to temperature fluctuations in the incubators used.

Table 6.3.14 Daf-c of *m65* and *tm1236* at 25°C trial 1

Genotype	% Dauers	% Dead Eggs	% L1 & L2	% Late larval & Adults	N
N2	0	0	4.0	96.0	151
daf-16(mgDf50)	0	0.5	5.4	94.1	204
daf-2(m65)/ qC1	23.0	0	1.0	76.0	183
daf-16(0); daf-2(m65)	0	31.3	44.8	23.9	96
daf-2(tm1236)/ qC1	26.0	0	2.0	72.0	154
daf-16(0); daf-2(tm1236)	0	0	6.84	93.16	117

Table 6.3.15 Daf-c of *m65* and *tm1236* at 25°C trial 2

Genotype	% Dauers	% Dead Eggs	% L1 & L2	% Late larval & Adults	N
N2	0	0	7.4	92.6	394
daf-16(mgDf50)	0	0	2.0	98.0	197
daf-2(m65)/ qC1	12.9	0	11.8	75.3	183
daf-16(0); daf-2(m65)	0	42.2	22.3	35.5	121
daf-2(tm1236)/ qC1	15.8	0	7.2	77.0	139
daf-16(0); daf-2(tm1236)	0	0.8	12.8	86.4	186

Dauer formation assays in *m65* and *tm1236* single mutants and as *daf-16*; *daf-2* double mutants. In the fist trial both the *m65/ qC1* and *tm1236/ qC1* segregate ~25% dauers and ~75% adults (of which ~25% are Dpy Glp). In the second trial ratio is masked by some early larval arrest. In both trial the *daf-16(mgDf50); daf-2(m65)* double mutant has high levels of embryonic and early larval arrest. Whereas, the majority of *daf-16(mgDf50); daf-2(tm1236)* double mutants proceed onto adulthood, suggesting that *tm1236* is almost fully suppressed by *daf-16(0)* and has more DAF-2 activity than *m65*.

The *tm1236* allele must be kinase-dead, since its deletion removes a section of exon 14 that encodes the catalytic and activation loops of the DAF-2 receptor. However, it is clear from the data above that *tm1236* is not a null allele of *daf-2*, as it is almost fully suppressed by *daf-16(0)*. This allele therefore retains some DAF-2 activity.

The *daf-16(0); daf-2(m65)* double mutant displays a 30-40% embryonic lethal phenotype at 25°C, indicating that DAF-2 has a DAF-16 independent function during

embryogenesis. One possibility is that the *tm1236* allele is actually a null allele, and the *m65* protein product has an antimorphic activity. I decided to test this by using RNAi of the *daf-2* gene on both double mutants to see if reducing transcript levels affected their phenotype. If the *m65* allele is antimorphic, then RNAi of *daf-2* might reduce its mutant severity. If it is a null then RNAi of *daf-2* should not increase the severity of its phenotype.

Table 6.3.16 RNAi of *tm1236* and *m65* at 25°C trial 1

Genotype	RNAi treatment	% Dauers	% Dead Eggs	% L1 & L2	% Late larval & Adults	N
daf-16(0); daf-2(m65)	L4440	0	92.4	6.8	0.8	132
daf-16(0); daf-2(m65)	*daf-2*	0	25.6	23.3	51.1	43
daf-16(0); daf-2(tm1236)	L4440	0	0	3.8	96.2	292
daf-16(0); daf-2(tm1236)	*daf-2*	0	0	4.5	95.5	286

Table 6.3.17 RNAi of *tm1236* and *m65* at 25°C trial 2

Genotype	RNAi treatment	% Dauers	% Dead Eggs	% L1 & L2	% Late larval & Adults	N
daf-16(0); daf-2(m65)	L4440	0	83.1	14.7	2.2	177
daf-16(0); daf-2(m65)	*daf-2*	0	78.1	20.7	1.2	169
daf-16(0); daf-2(tm1236)	L4440	0	0.9	1.8	97.3	114
daf-16(0); daf-2(tm1236)	*daf-2*	0	0	5.1	94.9	118

RNAi of the *daf-2* gene in *daf-16; daf-2* double mutant backgrounds. Neither the *daf-16(mgDf50); daf-2(m65)* double mutant or the *daf-16(mgDf50); daf-2(tm1236)* double mutant is affected by further reduction in levels of DAF-2 by using RNAi to knockdown *daf-2* transcript levels.

In the first trial at this temperature over 50% of RNAi treated *daf-16(0); daf-2(m65)* progeny reached adulthood, whereas only 0.8% of controls reached adulthood. In the second trial only 1.2% and 2.2% of progeny from *daf-2* RNAi and control worms

reached adulthood, respectively. I conclude that RNAi of the *daf-2* gene in *daf-16(0)*; *daf-2(m65)* and *daf-16(0)*; *daf-2(tm1236)* at 25°C has no effect on the severity of the phenotypes of these double mutants. This implies that *m65* is the best candidate molecular null allele of daf-2, and is not antimorphic.

6.4 Discussion

In this chapter I attempted to experimentally test some of the hypotheses about the role of DAF-2 signalling flux in determining the class 1 and 2 phenotypes that were generated in the previous chapter. Specifically, I asked whether the *e1369* allele is a regulatory mutant that affects *daf-2* transcript level. I evaluated the potential of two different DAF-2 antibodies in detecting processing defects in *daf-2* mutants that might leads to reduced levels of the receptor at the cell surface. I have also used RNAi of the *daf-2* gene in class 1 mutants to see if they develop class 2 phenotypes in response to reduced *daf-2* transcript levels. I also probed the potential role of ligand-receptor interaction in determining the class 1 and class 2 phenotypes by constructing *daf-2*; *daf-28* double mutants. Finally, I compared a nonsense allele and a deletion allele in order to identify a null allele of *daf-2*.

6.4.1 *daf-2* transcript levels in the *e1369* allele

In the previous chapter I identified the molecular lesion in several alleles of *daf-2*. However, I failed to find a mutation in the either coding sequence of the *e1369* allele or the UTRs of its transcript. I confirmed that this allele was actually a *daf-2* mutant by performing complementation tests with two alleles of *daf-2* whose mutations had been identified. The lack of mutation in the coding sequence suggests that the lesion in *e1369* affects a regulatory region of the *daf-2* gene. This might lead to reduced expression of *daf-2* mRNA and in turn reduced levels of DAF-2 receptor in the worm. This is consistent with my hypothesis that class 1 alleles of *daf-2* have reduced wild-type signalling flux. I sequenced ~5.2Kb of genomic DNA upstream of the start codon as well as the entire 5.5kb first intron of the *daf-2* gene from *e1369* worms, which is where one would expect to find most regulatory elements, but failed to find any mutations.

Given that the *daf-2* gene spans ~33Kb and the 5' and 3' intergenic spaces occupy a further ~20Kb (Chapter 4), rather than carry on sequencing I decided to directly look for reduced transcription of the *daf-2* gene in *e1369* by using QPCR. The *e1369* allele has the most severe Daf-c phenotype of all ts *daf-2* mutants, even though it is class 1, and forms 55% dauers at 15°C. This phenotype makes it very difficult to grow stage-

specific cultures of this mutant for the RNA samples needed for QPCR, especially since at high population density 100% of *e1369* worms arrest as dauers. To get around this I measured the *daf-2* transcript levels in *daf-16(0)*; *daf-2(e1369)* double mutants, which are Daf-d and can easily be grown in large stage-synchronous cultures. *daf-16(0)* worms were used as the control as these worms have a wild-type *daf-2* gene.

One potential concern about using *daf-16(0)* mutants is the possibility that DAF-16 regulates transcription of *daf-2*. This is certainly true for the orthologues of these genes in *D. melanogaster*, where dFOXO up-regulates the expression of the *dinr* gene (Puig *et al.*, 2003). If DAF-16 does regulate *daf-2* gene expression then using the *daf-16(0)* null allele may mask any DAF-16 dependent differences in gene expression between the *e1369* allele and wild-type *daf-2*. However, the microarray data from McElwee *et al* (2004) does not show a significant difference in the expression of the *daf-2* gene between *daf-2* mutants and *daf-16*; *daf-2* mutants.

I performed a total of 10 trials of QPCR to measure the levels of *daf-2* transcript in *daf-16(0)*; *daf-2(e1369)*, *daf-16(0)*; *daf-2(m212)* and *daf-16(0)* mutants. I found that mean levels of *daf-2* transcript were lower in the *e1369* double mutants compared to the *daf-16(0)* control worms in only five of these trials, and in no case was this reduction significant (Table 6.3.1). In three trials the levels of transcript were similar between the two genotypes, while in two of the trials the levels of *daf-2* transcript was significantly higher in *e1369* double mutants than *daf-16(0)* controls.

It is difficult to interpret the results of these trials. Overall, the results suggest a trend towards a reduction in the levels of *daf-2* transcript in the *e1369* allele. One possibility is that the mutation in the *e1369* allele affects the transcription of the *daf-2* gene in a tissue-specific manner. It is known that *daf-2* regulates both dauer formation and longevity in a cell non-autonomous manor (Apfeld and Kenyon, 1998; Wolkow *et al.*, 2000). Therefore loss of DAF-2 activity in just a few cells can have large effects on the Daf-c and Age phenotypes of a worm. If the mutation in *e1369* abolishes or severely reduces the transcription of the *daf-2* gene in these key cells it may cause the severe Daf-c phenotype of this allele without causing a detectable decrease in the levels of *daf-2* transcript, which are measured from whole worm RNA preparations. Currently there are no methods available that allow tissue-specific RNA extraction in *C. elegans*. Possibly, one could probe this hypothesis by looking at tissue specific effects on DAF-16 target gene expression using GFP.

6.4.2 DAF-2 antibodies

In the previous chapter I predicted that several alleles of *daf-2* have mutations that might cause processing defects in the DAF-2 proreceptor leading to a reduction in the level of mature receptor at the cell surface. I was interested in seeing if it was possible to detect the processing defects using DAF-2 antibodies. As a first step towards this I evaluated the potential of two different DAF-2 antibodies at detecting the DAF-2 protein using western blotting.

The first was a commercially available DAF-2 antibody from Santa Cruz Biotechnology Inc. This antibody was raised against a peptide that maps to the N-terminus of the receptor. The second was anti-peptide antibody that mapped to the C-terminus. I deliberately targeted the C-terminal extension of the DAF-2 protein for two reasons. Firstly, I was interested in whether this feature is cleaved off from the receptor. It has been shown that the C-terminal extension of dInR is cleaved off in a cell-specific manner (Fernandez *et al.*, 1995). Secondly, at this time point I had identified two nonsense mutants, *m65* and *m631*, of *daf-2* and informed of a third *m646* (K. Kimura and G. Ruvkun, personal communication). All three of nonsense mutations are predicted to truncate the DAF-2 receptor before the C-terminal extension. Therefore protein samples from these alleles might be useful controls in western blotting, as a C-terminal antibody should not bind anything.

Unfortunately, I was unable to detect any bands with the anti-DAF-2 antibody under conditions in which the MH27 antibody gave a positive signal. The cN-17 antibody detected several bands in both N2 and *daf-16(mgDf50)*; *daf-2(m65)* protein samples. However these were same size in both protein samples, which suggested that this antibody was not binding to DAF-2, as the protein from *daf-16(0)*; *daf-2(m65)* worms should run at a lower molecular weight as it is ~400aa shorter than wild-type. It may be that I did not use a high enough concentration of these antibodies, as I only went as high as 1:500 dilutions. It would be beneficial to try taking the antibody concentration as high as 1:100. Another way to test these antibodies further is to try immunolocalization. Which if successful could be used to directly look at the subcellular localization of the DAF-2 receptor in mutants with predicted processing defects. Overall, antipeptide antibodies are less effective than antibodies raised against whole proteins or protein fragments, as the short peptide sequences used may not adopt

the same conformation it has in the native protein. Potentially, antibodies raised against the whole DAF-2 protein or one of its domains may be more effective.

With hindsight, the use of the MH27 monoclonal antibody as a positive control was a poor choice. This antibody is generally used in immunolocalization assays (Koppen *et al.*, 2001). The only published reference that mentions this antibody in the context of western blotting is Francis and Waterston (1991) in which its only mention is that it recognises a protein greater than 150KDa in size. My western blots of N2 protein samples with MH27 generated multiple bands (Figure 6.4 and 6.5), some of which were in the correct size range as indicated in Francis and Waterston (1991). The antigen of the MH27 antibody has been recently identified (Koppen *et al.*, 2001). The MH27 antibody in immunolocalization assays binds to the AJM-1 protein, a component of the apical junction.

The *ajm-1* gene is predicted to have five different splice isoforms, which may account for some of the multiple bands seen in the N2 sample. However, I was able to confirm that the MH27 antibody was binding to the AJM-1 isoforms by using the SU159 strain. This strain of *C. elegans* is homozygous for a deletion allele of *ajm-1*, which is lethal. This strain is rescued by an extra-chromosomal array containing exons 2-8 of the *ajm-1* gene fused to GFP (Koppen *et al.*, 2001). A western blot of protein from this strain has a bright band that corresponds to the fusion protein and is missing two bands compared to N2 protein samples, which correspond to the two largest isoforms of AJM-1 (Figure 6.5).

My initial attempts at western blotting with both of these antibodies and my control antibody MH27 were complicated by the fact that my secondary antibodies seemed to binding worm proteins (Figure 6.2). These secondary antibodies were polyclonal so I decided to try using monoclonal anti-Goat and anti-Rabbit antibodies for my DAF-2 primary antibodies and an anti-Mouse Fc-specific antibody for MH27. These new secondary antibodies did not appear to react with any worm proteins (Figure 6.3), although the anti-Goat and anti-Rabbit monoclonals did seem to react with the molecular weight ladder. The binding of the secondary antibodies to *C. elegans* proteins suggests that the animals in which the antibodies were raised had parasitic nematode infections.

6.4.3 RNAi of the *daf-2* gene class 1 and 2 alleles

In the previous chapter I proposed that class 1 alleles of *daf-2* have mutations that lead to reduction of wild-type signalling flux from the receptor. Class 2 alleles have mutations that asymmetrically reduce signalling flux from the receptor, affecting the Shc/ MAP kinase pathway more than the PI3 kinase pathway. I have used RNAi of the *daf-2* gene to test the role of signalling flux in both classes of mutant as well as in *daf-2*; *daf-12* double mutants.

The critical result of this experiment is the effect of *daf-2* RNAi on *daf-2(m41)*; *daf-12(m20)* double mutants at 25°C. These mutants develop normally into fertile adults on L4440 control plates (Table 6.3.4 and 6.3.5). On *daf-2* RNAi, however, these worms arrest as sterile young adults in the same way that the *daf-2(e1370)*; *daf-12(m20)* mutants on *daf-2* RNAi at 20°C do (Table 6.3.2 and 6.3.3). This arrest phenotype may indicate that the *daf-2(41)*; *daf-12(m20)* double mutants have acquired a class 2 phenotype. If this were true it would mean that class 1 and 2 phenotypes correlate with absolute levels of DAF-2 activity and that my hypothesis about asymmetrical signalling flux is too simplistic.

It would be interesting to see if this holds true for other class 1 *daf-2*; *daf-12* double mutants, as *m41* is an unusual allele in that affects a novel amino acid that is only present in the nematode members of the IR superfamily (Chapter 5). Possibly the effect of RNAi on *m41* is atypical of class 1 alleles. Although its severe Daf-c phenotype and the result above are consistent with idea that class 1 mutations generally reduce levels of the DAF-2 receptor at the cell surface without altering its signalling ability.

6.4.4 Interaction between *daf-2* and *daf-28*

One of the predictions from the previous chapter is that class 1 mutations remain responsive to ligand, whereas class 2 extracellular mutants have ligand-binding defects. I constructed double mutants of both class 1 and class 2 alleles with the *daf-28(sa191)* allele. The *daf-28(sa191)* allele encodes a mutant insulin-like peptide that has a Daf-c phenotype (Li *et al.*, 2003) and is believed to act as poisonous ligand. My expectation

was that class 1 mutants would be responsive to this ligand, whereas class 2 mutant would not.

At 15°C the *sa191* allele enhances the Daf-c phenotype of *m577* but not *m579* (Table 6.3.6 and 6.3.7). This supports our hypothesis that *m579* is ligand-binding defective but *m577* is not. This enhancement seems to increase with temperature, as the *daf-2(m577); daf-28(sa191)* double mutants Daf-c phenotype increases in severity and forms almost 100% dauers at both 20°C and 22.5°C (Table 6.3.8, 6.3.9, 6.3.10 and 6.3.11). Interestingly the *daf-2(m579); daf-28(sa191)* double mutants also start to develop a Daf-c phenotype that is more severe than either parental strain at 20°C. Potentially the predicted ligand-binding defect in the *m579* mutant receptor might make it more sensitive to the loss of activating ligands, further reducing the signalling flux from the receptor. Another possibility is that the alteration in the ligand-binding properties of this DAF-2 mutant actually makes this receptor more sensitive to antagonistic ligands, which reduce the signalling flux through the receptor.

At 25°C the *daf-2(m577); daf-28(sa191)* had ~60% embryonic lethality in the first trial but only ~13% in the second trial (Table 6.3.12 and 6.3.13). Though these trials need to be repeated several more times to get a more accurate reading for the embryonic lethality, as there seemed to be relatively more embryonic arrest in trial 1 for all strains, except for N2 and *daf-2(m579); daf-28(sa191)*, than in trial 2. There was large amount of embryonic lethality seen in the *daf-2(m579); daf-28(sa191)* double mutants in both trials at 25°C. This was far greater than any embryonic arrest seen in either parental strain. Potentially this is related the reasons outlined above.

Even though the *daf-2(m577); daf-28(sa191)* double mutants have some embryonic lethality at 25°C, which resembles a class 2 trait, they do not become Unc at this temperature (data not shown). This suggests that the class 1 alleles do not become class 2 with the introduction of *daf-28(sa191)*. It is more likely that the lethality can simply be explained by severe reduction in DAF-2 activity either by lack of activation by agonistic ligands or increased suppression of activity by antagonistic ligands, as a result of the poisonous effect of the *daf-28(sa191)* allele.

6.4.5 Phenotypic comparison of nonsense and deletion alleles of *daf-2*

A comparison of the *m65* nonsense allele and the *tm1236* deletion allele shows that *m65* is still the best candidate for a null allele of *daf-2*. One possible explanation for the full suppression of the *tm1236* allele by *daf-16(mgDf50)*, even though it is missing more of the kinase domain than the *m65* allele, is the presence of the C-terminal extension in this allele. The properties of *tm1236* imply that the retention of the C-terminal extension in this allele allows the receptor to retain some residual activity. The C-terminal extension of DAF-2 contains four putative SH2-binding sites (Kimura *et al.*, 1997). These sites may allow the DAF-2 receptor to interact directly with substrates such as AAP-1 (Wolkow *et al.*, 2002). Even in the absence of ligand stimulated autophosphorylation the C-terminal extension of the *tm1236* mutant receptor may have enough residual basal phosphorylation ability to recruit and activate substrates. Because the *m65* allele is predicted to make a mutant receptor that lacks the C-terminal extension it may not be able to recruit substrate proteins, which may explain the severity of its phenotype at 25°C. Thus, unexpectedly the characteristics of *tm1236* provide the first evidence for functional activity of the DAF-2 C-terminal extension.

It still remains to be seen if the *m65* allele is a true null or whether it simply has the greatest reduction of DAF-2 activity of the alleles that have been characterised. Ultimately the only way to be sure of the *daf-2* null phenotype is to isolate a deletion that removes the entire receptor.

Chapter 7

Analysis of the *ins-7* and *ins-35* genes encoding putative ligands for DAF-2

7.1 Introduction

The previous chapters of my thesis have dealt with the DAF-2 insulin/ insulin-like growth factor receptor. This final results chapter relates to the insulin-like peptides that might regulate the receptor. One of the original hypotheses at the beginning of this project was that different INS peptides might differentially affect DAF-2 function. If so, then one might expect that some mutations in the ligand-binding portion of the receptor differentially affect the binding of INS peptides. This in turn might differentially affect the phenotype of the mutant. Currently, relatively little is known about *ins* gene function. In this chapter I describe studies of two *ins* genes, *ins-7* and *ins-35*, which show opposite regulation by *daf-16*.

7.1.1 Insulin-like peptides in vertebrates

The insulin molecule is a small peptide hormone that is part of a structurally related superfamily of proteins that share a common 'insulin fold' (Murray-Rust *et al.*, 1992). In vertebrates the major role of insulin is to regulate blood glucose concentration. However, it also stimulates cell growth and differentiation and promotes substrate storage in various tissues (Saltiel and Kahn, 2001). Insulin is not the only member of this superfamily found in vertebrates; additional members include IGF-1 and IGF-II, which have important roles in normal growth and development as well as in maintenance of differentiated tissues (Le Roith and Butler, 1999). Other members of the insulin superfamily in vertebrates include the relaxins, which in humans comprise of relaxin-1, -2, -3 and the insulin-like peptides (INSL), INSL3, INSL4, INSL5 and INSL6 (Wilkinson *et al.*, 2005). Relaxin-1, -2 and INSL3 have been shown to have a role in pregnancy and reproduction, while the function of other members of the relaxin family remains unclear.

The precursor proteins of these peptides in vertebrates have a very similar domain arrangement of a signal sequence followed by the B chain then the C peptide and the A chain (Sherwood, 2004). The IGF-1 and IGF-II precursors have two additional domains that are C-terminal of the A chain known as the D and E domains (Smit *et al.*, 1998). Even though the overall sequence similarity between the different peptides is low they

all contain six cysteine residues at invariant positions across the protein sequence, two in the B chain and four in the A chain. These cysteines are involved one intra-chain and two inter-chain disulphide bonds, which gives rise to the 'insulin fold' that structurally relates these different proteins. The C peptide of the insulin, relaxin-2, -3 and INSL3 precursors is cleaved during processing to give rise to the mature hormones which are dimers of the A and B chains. It is predicted that other relaxins also have their C peptide removed during maturation (Wilkinson *et al.*, 2005). In contrast the IGF-1 and IGF-II proteins retain their C peptides such that the mature hormone is formed from a single polypeptide.

7.1.2 Insulin-like peptides in invertebrates

Insulin-like peptides have now been identified in several different invertebrate species including *C. elegans* (reviewed in (Pierce *et al.*, 2001)), *Drosophila melanogaster* (Brogiolo *et al.*, 2001), *Anopheles gambiae* (Krieger *et al.*, 2004), *Locusta migratoria* (Lagueux *et al.*, 1990), three Lepidopterans including *Bombyx mori*, *Samia cynthia ricini* and *Agrius convoluvuli* (reviewed in (Yoshida *et al.*, 1998)) and in the mollusc *Lymnaea stagnalis* (reviewed in (Smit *et al.*, 1998)). A general feature of insulin-like peptides in invertebrates is that they are present in multiple copies encoded by multi-gene families. For example five different molluscan insulin-related peptides (MIPs) have been identified in *L. stagnalis* (Smit *et al.*, 1998) and seven *D. melanogaster* and *A. gambiae* insulin-like peptides have been found (DILPs and AgamILPs respectively) (Brogiolo *et al.*, 2001; Krieger *et al.*, 2004). In *B. mori* there are at least thirty-two bombyxin genes encoding insulin-like peptides (Yoshida *et al.*, 1998). All of the insulin-like peptides characterised in the insects have the canonical six cysteine residues forming the two inter-chain and single intra-chain disulphide bonds and a C peptide that is predicted to be cleaved. The five MIPs also have a C peptide that is predicted to be cleaved as well the six canonical cysteines. However they all possess an additional cysteine residue in both the A and B chains of the proteins that are expected to form a third inter-chain disulphide bond (Smit *et al.*, 1998).

In *C. elegans* there are 39 genes that encode insulin-like peptides, *ins-1* to *ins-37* (Duret *et al.*, 1998; Gregoire *et al.*, 1998; Kawano *et al.*, 2000; Kawano *et al.*, 2003; Pierce *et al.*, 2001), *ins-38* (M. Costa, personal communication) and *daf-28* (Li *et al.*,

2003). These insulin-like peptides are highly divergent from the other invertebrate insulin-like peptides. Only two of these genes, *ins-1* and *ins-18*, are predicted to have C peptides that are cleaved and form dimers like vertebrate insulin. The rest of the insulin-like peptides do not have a C peptide and are expected to remain single polypeptides like IGF-1 and IGF-II, although they do not possess the additional D and E domains of these proteins. The insulin-like peptides can be classified into four structural groups type γ, β and α and multiple chain (Duret *et al.*, 1998; Pierce *et al.*, 2001). The genes *ins-11*, *-12*, *-13*, *-14*, *-15*, *-16*, *-17*, *-18*, *-19*, *-32*, *-37* and *-38* encode the γ-type insulin-like peptides, which have the canonical six cysteines and are predicted to have the typical disulphide bond pattern of vertebrate insulins. The genes *ins-1*, *-2*, *-3*, *-4*, *-5*, *-6*, *-7*, *-8*, *-9*, *-10* and *daf-28* encode the β-type peptides. These peptides have two additional cysteine residues, one in the B chain and one in the A chain, both are C-terminal of the canonical cysteines in each chain. These cysteines are predicted to form an additional disulphide linking the A and B chain, similar to the additional disulphide bond seen in the MIPs, however their extra cysteines are N-terminal of the canonical cysteines of each chain (Smit *et al.*, 1998). The genes *ins-20*, *-21*, *-22*, *-23*, *-24*, *-25*, *-26*, *-27*, *-28*, *-29*, *-30*, *-33*, *-34*, *-35* and *-36* encode the α-type peptides. These peptides also have the additional cysteines involved in the third inter-chain disulphide bond in the β-type peptides. However they lack two of the canonical cysteines in the A chain responsible for the intra-chain bond. The *ins-31* gene encodes a novel multi-chain protein with the following domain arrangement; signal sequence - B chain - A chain - B chain - A chain - B chain. It is unknown if this protein is cleaved into three separate insulin-like peptides.

7.1.3 The role of insulin-like peptides in *C. elegans*

Given the large number of insulin-like peptides in *C. elegans* it is unsurprising that there is only limited information on their functional role and whether they interact with the IIS pathway. It has been shown that overexpression of *ins-1* and *ins-18*, as well as human insulin, results in a weak Daf-c phenotype for wild type worms (26°C) and increases the severity of the Daf-c phenotype in *daf-2(e1365)* worms (20°C). This suggests that these genes antagonise *daf-2* signalling (Pierce *et al.*, 2001). However, it has been shown that in cell culture chronic exposure to insulin leads to insulin

resistance (Garvey *et al.*, 1986). Thus, INS-1 and INS-18 may only be acting as DAF-2 antagonists at high ligand concentrations. Overexpression of *ins-1* and human insulin also increases the mean lifespan of wild type worms over controls, while a null allele of *ins-1* has no effect on lifespan or dauer formation. Overexpression of *ins-9*, *-22* and of an *ins-19/ -31* double construct has no effect on dauer formation in either wild type or *daf-2(e1365)* worms. However, overexpression of both *ins-9* and the *ins-19/ -31* double construct cause a significant increase in the amount of embryonic/ L1 arrest in *daf-2(e1365)* worms (26°C), a phenotype not normally associated with this class 1 allele of *daf-2*. Potentially these three genes also antagonise *daf-2* signalling but only function during embryogenesis or early larval development. Expression of *ins-19* and *-31* but not *ins-9* gene expression has been observed during embryogenesis (Pierce *et al.*, 2001).

The *daf-28* gene encodes the only insulin-like peptide in *C. elegans* in which a mutant has been shown to have a phenotype. The *sa191* allele of *daf-28* has a Daf-c phenotype at all temperatures, although the dauers recover within 2 hours, and increases lifespan by 12 -13% compared to wild type controls (Malone *et al.*, 1996). This allele behaves in a semi-dominant manner and has a missense mutation that causes a R37C substitution affecting the predicted proteolytic cleavage site between the signal sequence and the B chain (Li *et al.*, 2003). The mutant protein is believed to be poisonous to wild-type DAF-28 function and other INS proteins. Overexpression of *ins-4* and *ins-6*, which encode proteins with the highest sequence similarity to DAF-28, can rescue the Daf-c phenotype of *daf-28(sa191)*. Interestingly the Daf-c phenotype of *daf-28(sa191)* is only partially suppressed by hypomorphic alleles of *daf-16* implying that wild-type DAF-28, or an INS protein that the mutant poisons, interacts not only with the DAF-2/ DAF-16 pathway but some other parallel pathway that regulates dauer formation. This is unlikely to be the DAF-7 TGF-β pathway as Daf-d mutants in both *daf-3* and *daf-5* fail to suppress dauer formation in *daf-28(sa191)* (Malone *et al.*, 1996).

The role of several *ins* genes has been probed through the use of RNA-mediated interference (RNAi). RNAi of *ins-18* has been shown to extend lifespan in wild type worms (Kawano *et al.*, 2000), although this result is controversial, as a second study has shown that RNAi of *ins-18* does not extend lifespan in the RNAi hypersensitive strain *rrf-3(pk1426)* and leads to a small suppression of the lifespan extension seen in *daf-2(mu150)* (Murphy *et al.*, 2003). In this same study RNAi of both *ins-1* and *ins-7* increased the lifespan of *rrf-3(pk1426)* by 21.9% and 33.7% respectively, while RNAi of *ins-21* and *ins-22* had no effect. Similarly RNAi of *daf-28* has no effect on either

lifespan or dauer formation in *rrf-3(pk1426)* worms (Li *et al.*, 2003). The extension of lifespan by *ins-1* RNAi in *rrf-3(pk1426)* worms (Murphy *et al.*, 2003) is hard to reconcile with to the lifespan extension seen in wild type worms by *ins-1* overexpression (Pierce *et al.*, 2001). A possible explanation for these results might be that INS-1 is an agonist of DAF-2, rather than an antagonist as suggested by Pierce *et al* (2001). This would explain why RNAi of *ins-1* leads to lifespan extension. This would also imply that overexpression of *ins-1* leads to down-regulation of the *daf-2*/ IIS pathway in a similar way that chronic exposure to insulin cause cultured cells to become insulin-resistant (Garvey *et al.*, 1986). The suppression of lifespan by RNAi of *ins-18* in a *daf-2* mutant background (Murphy *et al.*, 2003) and an increase in the Daf-c phenotype of *daf-2(e1365)* at 20°C by *ins-18* overexpression suggest that INS-18 is a genuine antagonist of DAF-2, casting further doubt on the claim by Kawano *et al* (2000) that *ins-18* RNAi causes lifespan extension.

A key question concerning the INS peptides is why are there so many of them? There are several different possibilities. Firstly, the different INS proteins have different spatial and temporal expression (Li *et al.*, 2003; Pierce *et al.*, 2001). This raises the possibility of that some of these peptides have autocrine or paracrine roles. A second possibility is that not all of these peptides bind to the DAF-2 receptor. In vertebrates, members of the relaxin family have been shown to bind G-protein-coupled receptors (GPCR), (reviewed in Wilkinson *et al.*, 2005). The GPCR superfamily has undergone a large lineage-specific expansion in *C. elegans*, such that there are at least 1149 proteins belonging to this family (Fredriksson and Schioth, 2005). However, the relaxin interaction with GPCR proteins is dependent on a 'R-X-X-X-R-X-X-I' motif found in all members of the relaxin family. This motif is not present in any of the invertebrate insulin-like peptides (Wilkinson *et al.*, 2005), which suggests that if the INS peptides do interact with GPCR proteins they do so in a novel manner.

There is a third, less likely possibility. The *C. elegans* genome has 56 genes that encode members of a novel protein family. These proteins partly resemble the α-subunit of the DAF-2 receptor (Dlakic, 2002). However, these proteins, whilst having two full L-domains, have a truncated CR domain. These proteins also lack all three fibronectin type III domains, but do have a transmembrane domain. Dlakic (2002) suggests that these proteins may bind some INS peptides and that they in turn modulate the activity of DAF-2 through some novel interaction. However, it should be noted that the probable second contact site for the binding of insulin-like peptides in IR superfamily proteins is

located in the Fn1 domain, as shown by *in vitro* mutagenesis of *h*IR and *h*IGF-1R (Mynarcik *et al.*, 1997). This raises some doubt over whether the INS peptides could actually bind to these α-subunit-like proteins.

7.1.4 Microarray analysis of *ins* gene expression

Recently microarray data from our lab has shown that of the 23 *ins* genes represented on the Affymetrix™ *C. elegans* whole genome oligonucleotide arrays eight show significant changes in gene expression when comparing *daf-2* mutants to *daf-16*; *daf-2* double mutants (McElwee *et al.*, 2004) (Table 7.1.1). Of these, two genes, *ins-7* and *ins-35*, show the greatest change in expression levels. *ins-7* being down-regulated by at least ~75% in *daf-2* mutants compared to *daf-16*; *daf-2* double mutants, while *ins-35* is up-regulated ~24-fold in *daf-2* mutants. The down-regulation of *ins-7* in *daf-2* mutants is consistent with the view that INS-7 is a DAF-2 agonist, based on the extension of lifespan by *ins-7* RNAi in *rrf-3(pk1426)* worms (Murphy *et al.*, 2003). The up-regulation of *ins-35* in *daf-2* mutants suggests that INS-35 may antagonise DAF-2 function and therefore function as a longevity hormone.

In this chapter I further characterise changes in the expression of *ins-7* and *ins-35* and test their role in lifespan using RNA-mediated interference (RNAi). I have used quantitative PCR (QPCR) to measure the relative transcript levels of both *ins-7* and *ins-35* in *daf-2* mutants compared to *daf-16*; *daf-2* mutants. To test the role of *ins-7* and *ins-35* on lifespan I used a lactose-induced RNAi protocol that is more penetrant than isopropylthio-β-D-galactoside (IPTG) induced RNAi (J. J. McElwee, D. S. Patel and D. Gems, unpublished results) in several different genetic backgrounds.

Table 7.1.1 Effects of reduced IIS on *ins* gene expression

Gene	Structural Type	Ratio of *daf-2/ daf-16; daf-2* gene expression	Log2 Ratio	*p*-value
ins-21	α	0.899	-0.153	0.97856
ins-22	α	**3.658**	**1.871**	**0.00134**
ins-23	α	**2.433**	**1.283**	**0.01026**
ins-24[a]	α	**4.084**	**2.030**	**0.00893**
ins-24[a]	α	1.084	0.117	0.79351
ins-26	α	1.901	0.927	0.47522
ins-30	α	**2.518**	**1.332**	**0.00938**
ins-33	α	**3.371**	**1.753**	**0.00039**
ins-34	α	0.886	-0.175	0.81275
ins-35	α	**24.658**	**4.624**	**$9.4e^{-14}$**
ins-1	β	1.228	0.296	0.34290
ins-2	β	1.079	0.110	0.99168
ins-3	β	0.588	-0.766	0.19742
ins-4	β	0.742	-0.431	0.50619
ins-5	β	**1.480**	**0.566**	**0.06508**
ins-6	β	1.210	0.275	0.35704
ins-7	β	**0.251**	**-1.997**	**$4.7e^{-7}$**
daf-28	β	1.025	0.035	0.66744
ins-11	γ	**3.091**	**1.628**	**0.00692**
ins-17	γ	1.131	0.177	0.54407
ins-18	γ	1.084	0.117	0.81559
ins-32	γ	0.694	-0.528	0.23891
ins-37	γ	0.927	-0.109	0.95590
ins-31	Multi-chain	1.494	0.579	0.08119

Ratio of the expression levels of the 23 *ins* genes present on the Affymetrix™ array in *glp-4; daf-2* double mutants compared to *daf-16 glp-4; daf-2* triple mutants. Genes listed in **bold** have significantly altered expression levels between the two strains (p-value < 0.05). [a] There are two different probe sets for *ins-24* on the Affymetrix™ array, only one of which gives a significant change in gene expression. **Data from McElwee *et al* (2004)**

7.2 Material and Methods

7.2.1 Strains

The following strains were used for quantitative analysis of the transcript levels of *ins-7* and *ins-35*: GA133 *daf-16(mgDf50) glp-4(bn2) I; daf-2(m577) III* and GA134 *glp-4(bn2) I; daf-2(m577) III,* (McElwee *et al.*, 2004). The following strains were used for RNAi lifespans: DR1567 *daf-2(m577) III* (Gems *et al.*, 1998); GA134 *glp-4(bn2ts) I; daf-2(m577) III*; GA303 *rrf-3(pk1426) II; daf-2(m577) III;* SS104 *glp-4(bn2ts) I* (Beanan and Strome, 1992); NL2099 *rrf-3(pk1426) II* (Simmer *et al.*, 2002) and N2 CGC male stock (Gems and Riddle, 2000).

The GA303 strain *rrf-3(pk1426)*; *daf-2(m577)* was constructed by crossing *daf-2(m577)* males with *rrf-3(pk1426)* hermaphrodites at 15°C. The hermaphrodite progeny of this cross were raised until they were gravid adults then shifted to fresh plates at 25°C. These animals were allowed to lay eggs for 24 hours before being removed from the plates. The progeny were kept at 25°C for three days, after which they were examined for the presence of dauers. Approximately 50 dauers, homozygous for *daf-2(m577)*, were picked to fresh plates and allowed to recover at 15°C to the L4 stage. Once the animals were at L4 they were plated singly at 25°C and left for three days. The *rrf-3(pk1426)* allele causes a drastic reduction in fecundity of the hermaphrodites at 25°C (Sijen *et al.*, 2001). The plates at 25°C were checked for the number of progeny (all dauers) and those with <10 were taken to be homozygous for *rrf-3(pk1426)* and shifted down to 15°C to allow recovery. Recovered L4s were then plated singly and shifted back to 25°C to confirm that they also had reduced fecundity at this temperature. After confirming the reduced fecundity phenotype, six dauers were picked from two plates and shifted back to 15°C to allow recovery.

7.2.2 Quantitative PCR

The relative abundance of *ins-7* and *ins-35* transcripts in *glp-4(bn2); daf-2(m577)* mutant adults compared to *glp-4(bn2) daf-16(mgDf50); daf-2(m577)* mutant adults was

determined using QPCR. The RNA samples used in the first two QPCR experiments were the same as those prepared previously for microarray analysis by McElwee *et al* (2004) and included three samples of *glp-4(bn2); daf-2(m577)* RNA and three of *glp-4(bn2) daf-16(mgDf50); daf-2(m577)*. RNA samples for the third QPCR experiment were obtained using worms cultured at the same time as those in McElwee *et al* (2004) that had been kept stored in TRIzol (Invitrogen) at -80°C and not processed for the microarray work. These RNA samples included two *glp-4; daf-2* samples and one of *glp-4(bn2) daf-16(mgDf50); daf-2(m577)*.

The sequences of the primers used in the QPCR are listed in Table 7.2.1. The *ama-1* primers are the same as those described previously in chapter 6. The primers for the amplification of *ins-7* and *ins-35* transcripts were designed using the Primer3 WWW interface (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) (Rozen and Skaletsky, 2000). These primers do not span exon-exon boundaries, but do anneal to different exons in order to prevent the amplification of any contaminating genomic DNA within the RNA samples. Protocols for the RNA extraction, first strand cDNA synthesis and QPCR can be found in sections 3.2.1, 3.2.2 and 3.2.4 of the main materials and methods chapter.

Table 7.2.1 Primer pairs used in QPCR

Primer name	5'- Sequence - 3'
*ama-1*Fwd	CCCGGAGGAGATTAAACGCATG
*ama-1*Rev	CATCTTCCACGACGATCTATGA
*ins-7*Fwd	GAAGAGTCCCTGACGAGAAAAA
*ins-7*Rev	TCTTCACGGCAACATTTTGAT
*ins-35*Fwd	TATTCATTTGCCATCTGTGGAG
*ins-35*Rev	CTCTGGGCAGCAGATAAGTTTT

In the first two QPCR experiments analysis of the expression levels of *ins-7* and *ins-35* were carried on separate 96-well PCR plates due to the larger number of RNA samples used in these experiments. Each plate had the general layout seen in Table 7.2.2. In the third QPCR experiment both the *ins-7* and *ins-35* reactions were carried out on the same PCR plate (Table 7.2.3). All of the QPCR experiments were performed using the

relative standard curve method (Applied Biosystems user bulletin 2, http://docs.appliedbiosystems.com/pebiodocs/04303859.pdf).

The *ama-1* gene was used as the endogenous control in all experiments. The ratio of expression of *ama-1* in *daf-2* mutants compared to *daf-16*; *daf-2* mutants was 1.003 and the *p*-value was 0.976 in the microarray analysis of McElwee *et al* (2004). This implies that there is no significant change in expression level of this gene between the two genotypes, which supports the choice of this gene as an endogenous control.

Table 7.2.2 General QPCR plate layout

	1	2	3	4	5	6	7	8	9	10	11	12
A	NTC *ama-1*	NTC *ama-1*	NTC *ama-1*	2xN2 *ama-1*	2xN2 *ama-1*	2xN2 *ama-1*	1xN2 *ama-1*	1xN2 *ama-1*	1xN2 *ama-1*	1/2 N2 *ama-1*	1/2 N2 *ama-1*	1/2 N2 *ama-1*
B	1/4 N2 *ama-1*	1/4 N2 *ama-1*	1/4 N2 *ama-1*	1/8 N2 *ama-1*	1/8 N2 *ama-1*	1/8 N2 *ama-1*	1/16 N2 *ama-1*	1/16 N2 *ama-1*	1/16 N2 *ama-1*	1/32 N2 *ama-1*	1/32 N2 *ama-1*	1/32 N2 *ama-1*
C	1/64 N2 *ama-1*	1/64 N2 *ama-1*	1/64 N2 *ama-1*	1/128 N2 *ama-1*	1/128 N2 *ama-1*	1/128 N2 *ama-1*	*m577*/1 *ama-1*	*m577*/1 *ama-1*	*m577*/1 *ama-1*	*m577*/2 *ama-1*	*m577*/2 *ama-1*	*m577*/2 *ama-1*
D	*m577*/3 *ama-1*	*m577*/3 *ama-1*	*m577*/3 *ama-1*	*mgDf50; m577*/1 *ama-1*	*mgDf50; m577*/1 *ama-1*	*mgDf50; m577*/1 *ama-1*	*mgDf50; m577*/2 *ama-1*	*mgDf50; m577*/2 *ama-1*	*mgDf50; m577*/2 *ama-1*	*mgDf50; m577*/3 *ama-1*	*mgDf50; m577*/3 *ama-1*	*mgDf50; m577*/3 *ama-1*
E	NTC *ins-7/-35*	NTC *ins-7/-35*	NTC *ins-7/-35*	2xN2 *ins-7/-35*	2xN2 *ins-7/-35*	2xN2 *ins-7/-35*	1xN2 *ins-7/-35*	1xN2 *ins-7/-35*	1xN2 *ins-7/-35*	1/2 N2 *ins-7/-35*	1/2 N2 *ins-7/-35*	1/2 N2 *ins-7/-35*
F	1/4 N2 *ins-7/-35*	1/4 N2 *ins-7/-35*	1/4 N2 *ins-7/-35*	1/8 N2 *ins-7/-35*	1/8 N2 *ins-7/-35*	1/8 N2 *ins-7/-35*	1/16 N2 *ins-7/-35*	1/16 N2 *ins-7/-35*	1/16 N2 *ins-7/-35*	1/32 N2 *ins-7/-35*	1/32 N2 *ins-7/-35*	1/32 N2 *ins-7/-35*
G	1/64 N2 *ins-7/-35*	1/64 N2 *ins-7/-35*	1/64 N2 *ins-7/-35*	1/128 N2 *ins-7/-35*	1/128 N2 *ins-7/-35*	1/128 N2 *ins-7/-35*	*m577*/1 *ins-7/-35*	*m577*/1 *ins-7/-35*	*m577*/1 *ins-7/-35*	*m577*/2 *ins-7/-35*	*m577*/2 *ins-7/-35*	*m577*/2 *ins-7/-35*
H	*m577*/3 *ins-7/-35*	*m577*/3 *ins-7/-35*	*m577*/3 *ins-7/-35*	*mgDf50; m577*/1 *ins-7/-35*	*mgDf50; m577*/1 *ins-7/-35*	*mgDf50; m577*/1 *ins-7/-35*	*mgDf50; m577*/2 *ins-7/-35*	*mgDf50; m577*/2 *ins-7/-35*	*mgDf50; m577*/2 *ins-7/-35*	*mgDf50; m577*/3 *ins-7/-35*	*mgDf50; m577*/3 *ins-7/-35*	*mgDf50; m577*/3 *ins-7/-35*

NTC= No template control

Table 7.2.3 Plate layout for third QPCR experiment

	1	2	3	4	5	6	7	8	9	10	11	12
A	NTC *ama-1*	NTC *ama-1*	NTC *ama-1*	1xN2 *ama-1*	1xN2 *ama-1*	1xN2 *ama-1*	1/2 N2 *ama-1*	1/2 N2 *ama-1*	1/2 N2 *ama-1*	1/4 N2 *ama-1*	1/4 N2 *ama-1*	1/4 N2 *ama-1*
B	1/16 N2 *ama-1*	1/16 N2 *ama-1*	1/16 N2 *ama-1*	1/64 N2 *ama-1*	1/64 N2 *ama-1*	1/64 N2 *ama-1*	1/128 N2 *ama-1*	1/128 N2 *ama-1*	1/128 N2 *ama-1*	NTC *ins-7*	NTC *ins-7*	NTC *ins-7*
C	1xN2 *ins-7*	1xN2 *ins-7*	1xN2 *ins-7*	1/2 N2 *ins-7*	1/2 N2 *ins-7*	1/2 N2 *ins-7*	1/4 N2 *ins-7*	1/4 N2 *ins-7*	1/4 N2 *ins-7*	1/16 N2 *ins-7*	1/16 N2 *ins-7*	1/16 N2 *ins-7*
D	1/64 N2 *ins-7*	1/64 N2 *ins-7*	1/64 N2 *ins-7*	1/128 N2 *ins-7*	1/128 N2 *ins-7*	1/128 N2 *ins-7*	NTC *ins-35*	NTC *ins-35*	NTC *ins-35*	4xN2 *ins-35*	4xN2 *ins-35*	4xN2 *ins-35*
E	2xN2 *ins-35*	2xN2 *ins-35*	2xN2 *ins-35*	1xN2 *ins-35*	1xN2 *ins-35*	1xN2 *ins-35*	1/2 N2 *ins-35*	1/2 N2 *ins-35*	1/2 N2 *ins-35*	1/4 N2 *ins-35*	1/4 N2 *ins-35*	1/4 N2 *ins-35*
F	1/16 N2 *ins-35*	1/16 N2 *ins-35*	1/16 N2 *ins-35*	1/64 N2 *ins-35*	1/64 N2 *ins-35*	1/64 N2 *ins-35*	1/128 N2 *ins-35*	1/128 N2 *ins-35*	1/128 N2 *ins-35*	*m577/1 ama-1*	*m577/1 ama-1*	*m577/1 ama-1*
G	*m577/2 ama-1*	*m577/2 ama-1*	*m577/2 ama-1*	*mgDf50; m577/1 ama-1*	*mgDf50; m577/1 ama-1*	*mgDf50; m577/1 ama-1*	*m577/1 ins-7*	*m577/1 ins-7*	*m577/1 ins-7*	*m577/2 ins-7*	*m577/2 ins-7*	*m577/2 ins-7*
H	*mgDf50; m577/1 ins-7*	*mgDf50; m577/1 ins-7*	*mgDf50; m577/1 ins-7*	*m577/1 ins-35*	*m577/1 ins-35*	*m577/1 ins-35*	*m577/2 ins-35*	*m577/2 ins-35*	*m577/2 ins-35*	*mgDf50; m577/1 ins-35*	*mgDf50; m577/1 ins-35*	*mgDf50; m577/1 ins-35*

NTC= No template control

7.2.3 RNAi protocols

Two different methods of induction of double-stranded RNA (dsRNA) expression in HT115(DE3) cells were employed in the RNAi experiments. These used either the IPTG method (Kamath *et al.*, 2000) or a lactose-based induction method based on a modification of a protocol originally described by Eric J. Lambie (personal communication).

For RNAi of *ins-7*, two different dsRNA expression plasmids were used. The first plasmid contained 370bp of genomic sequence that included the entire *ins-7* gene (kindly provided by J. J. McElwee). The second *ins-7* construct was from the Ahringer feeding library (Kamath *et al.*, 2003) and contains 1144bp of genomic DNA including the entire *ins-7* gene. This is the construct that has been shown to extend lifespan (Murphy *et al.*, 2003). To avoid confusion this will be referred to as the ZK1251.2 construct for the remainder of this chapter. RNAi of the *daf-2* gene was used as a positive control for lifespan extension, previously shown by Dillin *et al* (2002). I used the *daf-2* construct provided by Adam Antebi, described in chapter 6.

The presence of a second plasmid in the ZK1251.2 feeding strain was discovered accidentally while checking that the ZK1251.2 construct contained the correct length insert. PCR reactions were performed on cells from eight different bacterial colonies containing the ZK1251.2 construct grown on LB + ampicillin plates. The ZK1251.2 insert was amplified using primers that flanked the multiple cloning site of the L4440 vector (kindly provided by J. J. McElwee). Control reactions were carried on cells containing L4440, *ins-7* and *ins-35* constructs. When the PCR reactions were run on an agarose gel the control PCR reactions contained only a single band, whereas all eight of the ZK1251.2 reactions contained two distinct bands. One band was ~1.1kb in length and corresponded to the ZK1251.2 insert. The other band was smaller and was the same size as the band in the L4440 control reaction. This suggested that the HT115(DE3) cells containing the ZK1251.2 construct were also transfected with the empty L4440 plasmid.

For *ins-35* a new construct was prepared using the primers *ins-35*RNAiLeft and *ins-35*RNAiRight (Table 7.2.4) to amplify a 639bp fragment of genomic DNA, from N2 worms using a touchdown PCR method (Materials and Methods chapter). This fragment contains the entire *ins-35* gene, which was then ligated into the pCR®II-TOPO® vector

using the TOPO-TA cloning kit® and transformed into TOP10 *E. coli* cells (Invitrogen), following the manufacturer's instructions. Transformed cells were then grown in overnight culture and the plasmid DNA was extracted using a QIAprep miniprep kit (Qiagen). The isolated plasmid DNA was digested with *Xho*I and *Spe*I and ligated into the L4440 vector (Timmons and Fire, 1998), which had been linearized with the same restriction enzymes. The recombinant plasmid (pDP001) was transformed into chemically competent HT115(DE3) cells (Timmons *et al.*, 2001).

Table 7.2.4 Primer pairs for RNAi constructs

Primer name	5'-Sequence-3'
TJ1192	CCCAAGCTTCCATTTATGATGTGACATTGCTCGTGG
TJ1193	GGTTGAGATTGTTTCAAGCAATTGAACCC
TJ1194	CGGGGTACCCCAGATGCCACCAATAATTTTGG
TJ1195	TGCTCTAGAGAAGAAATTAAGGACAGCACTGTTTTCG
*ins-35*RNAiLeft	TGAAGCAAATATTCTTGGTAATCC
*ins-35*RNAiRight	CTCTGGGCAGCAGATAAGTTTT

In addition to the *ins* genes, several other genes that had shown significant up-regulation in the *daf-2* mutants in McElwee *et al* (2004) were also subjected to RNAi. These genes are listed in Table 7.2.5 and were part of larger RNAi screen performed in our lab, but in the present context served as further negative controls for the general effects of RNAi on lifespan.

Table 7.2.5 Genes up-regulated in *daf-2* mutants

Sequence Name	Gene Name/ Putative homologue?	Log2 Ratio of *daf-2*/ *daf-16*; *daf-2* gene expression	*p*-value
T22B7.1	*egl-13*	0.696	0.02766
C15F1.7	*sod-1*	0.807	0
H42K12.1	*pdk-1*	1.136	0.00029
R11A5.4	PEPCK	2.065	0.00029
C36A4.2	*cyp-25A2*	0.720	0.00048
C55F2.1	AICAR transfromylase	1.466	0.00175
C06A5.7	*tmd-1*	0.372	0.03636
W01A11.1	Epoxide hydrolase	1.631	0
K11E8.1	*unc-43*	0.239	0.0303
F21A3.2	Purple acid phosphatase	0.895	0.00027

Genes that show statistically significant up-regulation in *glp-4*; *daf-2* double mutants compared to *glp-4 daf-16*; *daf-2* triple mutants. These genes were part of a larger RNAi screen within our lab. **Data from McElwee *et al* (2004)**

7.2.4 Lifespan assays

The lifespan assays were carried out at 25°C using either M9 minimal media-lactose plates, NGM-lactose plates or IPTG plates, or at 20°C with IPTG plates. For the assays at 25°C all animals were raised at 15°C through at least two generations of RNAi before being picked at the L4 stage and shifted to the higher temperature. The majority of the assays had at least 100 worms to begin with, which were transferred every other day until the end of egg laying for non-sterile strains, after this worms were transferred as needed. Plates were scored for death every other day. Death was scored as failure of response to prodding with a worm pick. Worms that died from internal hatching of eggs or ruptured vulvas were censored. All dead worms were removed from the plates.

For the 20°C lifespans worms were also raised at 15°C till the third generation on RNAi. Worms were then picked at the L2 stage and transferred to 25°C until the first day of adulthood and shifted down to 20°C for the remainder of their lifespan. This

protocol is the same as that used by Murphy *et al* (2003) for their RNAi lifespans. Each assay had at least 100 worms, which were transferred every other day until the end of egg laying and then as needed. Death was scored as above. HT115(DE3) cells transformed with the L4440 vector were used as negative controls in all lifespans. Statistical analysis of lifespan data was carried out using the JMP IN® v5.1 statistics software package (SAS Institute Inc.). Survival curves were generated using the Kaplan-Meier method and the statistical significance of differences between the curves was determined using a log rank test. The *p*-value for each lifespan represents the probability of obtaining by chance a χ^2 value greater than the value obtained under the assumption that the survival curves between tested groups is the same.

7.3 Results

Previous microarray analysis of the gene expression changes between *daf-2* mutants and *daf-16*; *daf-2* double mutants had shown that the *ins-7* and *ins-35* genes were differentially regulated. *ins-7* is the most strongly down-regulated *ins* gene in *daf-2* mutants compared to *daf-16*; *daf-2*, while *ins-35* is the most up-regulated (McElwee *et al.*, 2004). This suggests that *ins-7* and *ins-35* may promote ageing and longevity, respectively. My aim in the following studies is to verify the role of *ins-7* in ageing, previously shown by Murphy *et al* (2003), and to test the role of *ins-35* in longevity assurance.

7.3.1 Relative abundance of *ins-7* and *ins-35* transcripts

I have used QPCR to confirm these changes and to estimate the relative transcript levels of each gene in *glp-4; daf-2* mutant adults compared to *glp-4(bn2) daf-16(mgDf50)*; *daf-2(m577)* mutant adults. The fold-changes for both *ins-7* and *ins-35* are listed in Table 7.3.1. The values were calculated using the relative standard curve method, in which the quantities of the target genes (*ins-7* and *ins-35*) in each RNA sample were normalized to the *ama-1* quantity for that sample, under the assumption that *ama-1* expression is invariant between genotypes and samples (see materials and methods, this chapter). The mean value of the normalized transcript level for the target genes was calculated for each genotype. The mean value of the target gene for the *daf-2* samples was divided by the mean value of the target gene in the *daf-16*; *daf-2* samples to give the estimated fold-change. Standard deviations were calculated as recommended in the Applied Biosystems user bulletin 2.

In the case of the first two QPCR experiments the values of the normalized quantities of *ins-7* and *ins-35* from the third RNA sample of *glp-4*; *daf-2* were not included in the calculations of the mean value or the standard deviation. This RNA sample consistently had a 30-fold higher level of *ama-1* than the other two *glp-4; daf-2* samples (data not shown) and therefore was treated as an anomalous outlier. To verify that this interpretation I prepared fresh RNA samples from both *glp-4; daf-2* and *glp-4 daf-16*; *daf-2* worms that had been cultured at the same time as the previous batch and

kept at -80°C. These RNA samples were then used in the third QPCR experiment (Table 7.2.3). The *ama-1* levels in the new *glp-4; daf-2* RNA samples were similar to the first two samples, confirming that the third sample was indeed an outlier.

Table 7.3.1 Relative transcript levels of *ins-7* and *ins-35*

QPCR plate	Relative expression level of *ins-7* (fold-change ± Std dev)		Relative expression level of *ins-35* (fold-change ± Std dev)	
	in *daf-2*	in *daf-16; daf-2*	in *daf-2*	in *daf-16; daf-2*
Plate 1	0.54 ± 0.28	1 ± 0.44	98.34 ± 0.75	1 ± 0.45
Plate 2	0.61 ± 0.59	1 ± 0.44	102.68 ± 0.71	1 ± 0.27
Plate 3	0.30 ± 0.26	1 ± 0.09	68.77 ± 0.20	1 ± 0.17

Relative expression levels of *ins-7* and *ins-35* in *glp-4; daf-2* double mutants compared to *glp-4 daf-16; daf-2* triple mutants. The expression level of *ins -7* is ~40% less in *glp-4; daf-2* than in *glp-4 daf-16; daf-2* in both plates 1 and 2 and ~70% less in plate 3. The expression levels of *ins-35* is ~100-fold higher in *glp-4; daf-2* than in *glp-4 daf-16; daf-2* in both plates 1 and 2 and ~70-fold higher in plate 3.

The fold-changes for *ins-7* gene expression in the first two QPCR experiments are 0.54 ± 0.28 and 0.61 ± 0.59, respectively (Table 7.3.1). Not only are these values higher than the estimated value of 0.251 from the microarray data (Table 7.1.1), but the standard deviation amongst the *daf-2* samples overlaps with that of the *daf-16; daf-2* samples, indicating that the expression levels are not significantly different. In the third QPCR experiment the standard deviations of the *daf-2* samples do not overlap with the standard deviation of the *daf-16; daf-2* sample, however in this experiment there is only one sample of *daf-16; daf-2* RNA, which probably explain the smaller standard deviation.

The fold-change in the level of *ins-35* expression is very large in all three QPCR experiments. In the first two experiments there is a ~100-fold increase in *ins-35* gene expression in the *daf-2* samples compared to *daf-16; daf-2* samples, while in the third experiment there is ~70-fold increase in the level of *ins-35*.

7.3.2 Testing the role of *ins-7* and *ins-35* on lifespan by RNAi

If the *ins-7* and *ins-35* genes promote ageing and longevity, respectively, then RNAi of these genes should increase and decrease lifespan, respectively. In total 10 different RNAi constructs were tested for their effects on lifespan in variety of genetic backgrounds. These constructs included two different *ins-7* plasmids, including one from the Ahringer feeding library (Kamath *et al.*, 2003), *ins-35* and *daf-2*, as well as the constructs listed in Table 7.2.5. Sequencing of the inserts from the C36A4.2 and C55F2.1 constructs (Table 7.2.5) from the Ahringer feeding library showed that these did not correspond to the correct genes. Therefore these clones were excluded from further analysis.

ins gene RNAi using M9 minimal media-lactose plates (Trial 1)

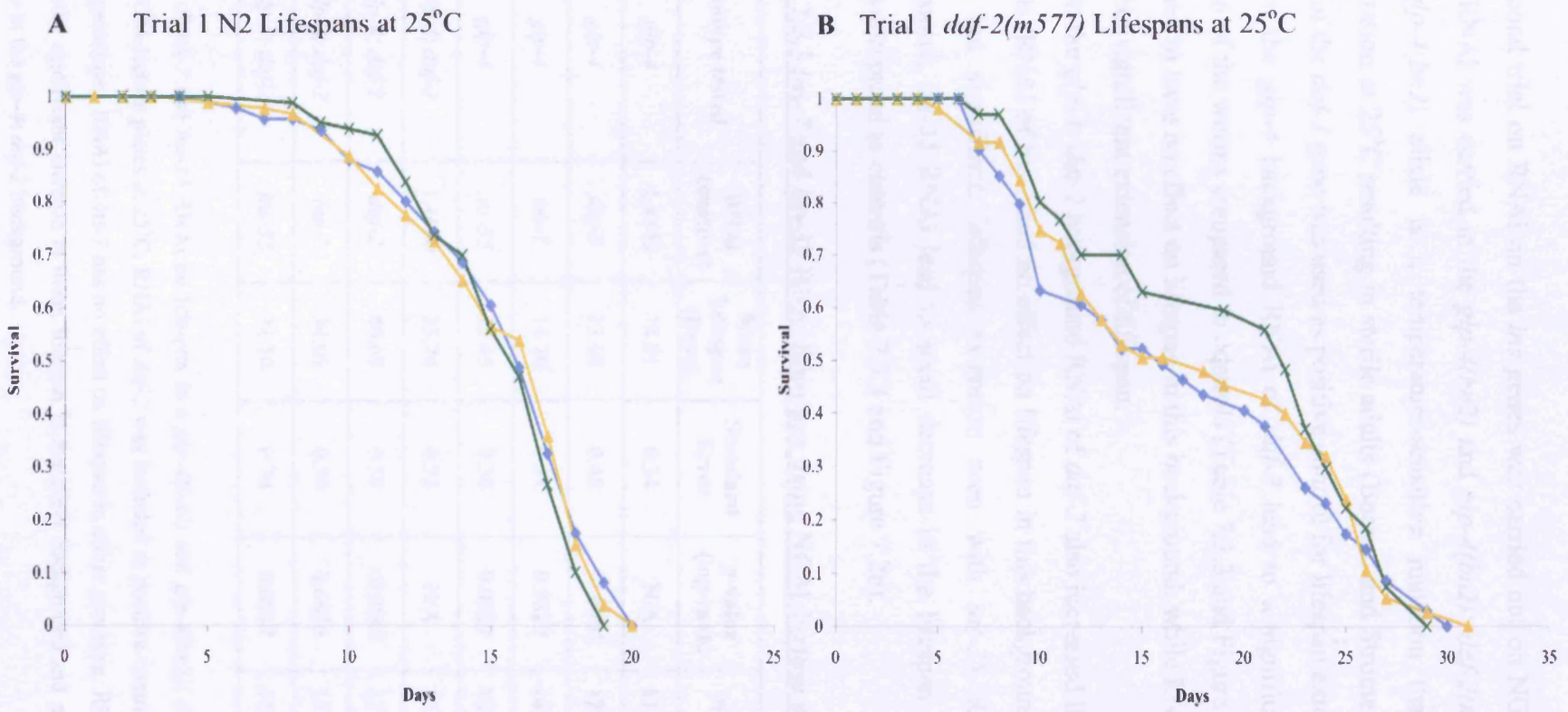
My first trial of *ins-7* and *ins-35* RNAi used M9 minimal media plates with lactose as the inducer of dsRNA expression. RNAi was carried on N2 and *daf-2(m577)* animals at 25°C. In the N2 background, RNAi of the *ins* genes had no significant effect on lifespan (Table 7.3.2 and Figure 7.1a). In the *daf-2(m577)* background RNAi of *ins-35* lead to a slight increase in the mean lifespan compared to controls, but there no significant effect on overall lifespan. The M9 plates seem to cause a significant amount of internal hatching in the *daf-2(m577)* background, but not in N2, which affects the shape of the lifespan curves in this background (Figure 7.1b). These plates also seem to partially reduce the extension lifespan of *daf-2(m577)* animals compared to the N2 animals. Potentially these plates were toxic to *daf-2* animals so all subsequent lifespans were carried on either standard IPTG plates or NGM-lactose plates.

Table 7.3.2 *ins-7* and *ins-35* RNAi lifespans on M9 media-lactose plates (Trial 1)

Genotype	RNAi construct	Mean Lifespan (days)	Standard Error	p-value (log-rank)	N	Censor	% Control Lifespan
N2	L4440	15.45	0.37	N/A	87	13	100
N2	*ins-7*	15.45	0.37	0.9964	81	19	99.99
N2	*ins-35*	15.46	0.31	0.2122	79	21	100.06
daf-2(m577)	L4440	17.07	1.27	N/A	36	64*	100
daf-2(m577)	*ins-7*	17.77	1.18	0.6969	28	72*	104.12
daf-2(m577)	*ins-35*	19.52	1.31	0.4297	40	65*	114.36

Effects of *ins-7* and *ins-35* RNAi on lifespan in a N2 and *daf-2(m577)* background using M9 minimal media plates at 25°C. Neither of these genes has any effect in the N2 background. Both lead to a slight increase in mean lifespan in the *m577* background, but this is not statistically significant. * The majority of these censored animals died from internal hatching

Figure 7.1 *ins-7* and *ins-35* RNAi lifespans at 25°C using M9 minimal media plates



RNAi of *ins-7* and *ins-35* in N2 background has no effect on lifespan at 25°C.

RNAi of *ins-7* and *ins-35* in *daf-2(m577)* background has no effect on lifespan at 25°C. However, a large proportion of *m577* animals on all 3 treatments died from internal hatching.

ins gene RNAi using NGM-lactose plates (Trial 2)

My second trial on RNAi on the *ins* genes was carried out on NGM-lactose plates at 25°C. RNAi was carried in the *glp-4(bn2)* and *glp-4(bn2); daf-2(m577)* backgrounds. The *glp-4(bn2)* allele is a temperature-sensitive mutation that blocks germline proliferation at 25°C resulting in sterile adults (Beanan and Strome, 1992). In this trial RNAi of the *daf-2* gene was used as positive control for lifespan extension.

In the *glp-4* background RNAi of *daf-2* lead to a significant increase in the lifespan of the worms compared to controls (Table 7.3.3 and Figure 7.2a). RNAi of *ins-7* seemed to have no effect on lifespan in this background, while RNAi of *ins-35* lead to small but significant extension of lifespan.

In the *glp-4; daf-2* background RNAi of *daf-2* also increased lifespan compared to controls. RNAi of *ins-7* had no effect on lifespan in this background. In contrast to the small, but significant, lifespan extension seen with *ins-35* RNAi in the *glp-4* background, *ins-35* RNAi lead to small decrease in the lifespan of the *glp-4; daf-2* worms compared to controls (Table 7.3.3 and Figure 7.2b).

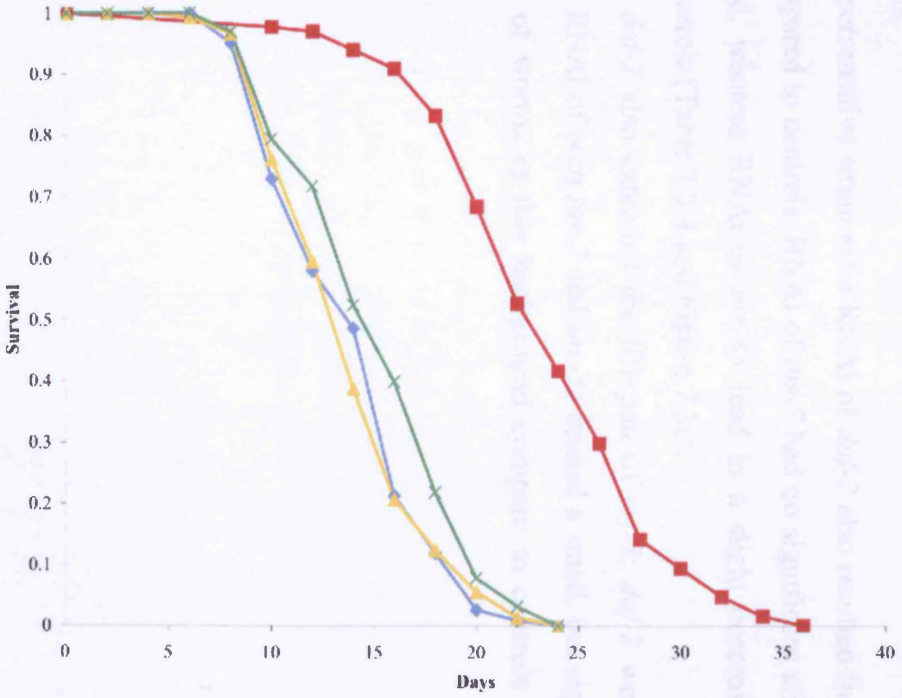
Table 7.3.3 *ins-7* and *ins-35* RNAi lifespans using NGM-lactose plates (Trial 2)

Genotype tested	RNAi construct	Mean Lifespan (Days)	Standard Error	*p*-value (log-rank)	N	Censor	% Control Lifespan
glp-4	L4440	14.21	0.34	N/A	119	7	100
glp-4	*daf-2*	23.68	0.48	<0.0001	128	3	166.57
glp-4	*ins-7*	14.20	0.31	0.9327	145	7	99.88
glp-4	*ins-35*	15.45	0.36	0.0087	129	4	108.71
glp-4; daf-2	L4440	33.74	0.72	N/A	134	10	100
glp-4; daf-2	*daf-2*	40.69	0.59	<0.0001	121	5	120.59
glp-4; daf-2	*ins-7*	34.09	0.59	0.6656	153	4	101.03
glp-4; daf-2	*ins-35*	31.10	0.74	0.0047	119	4	92.17

Effects of *ins-7* and *ins-35* RNAi on lifespan in a *glp-4(bn2)* and *glp-4(bn2); daf-2(m577)* background using NGM-lactose plates at 25°C. RNAi of *daf-2* was included as positive control for lifespan extension in both genotypes. RNAi of *ins-7* has no effect on lifespan in either genotype. RNAi of *ins-35* leads to a statistically significant increase in mean lifespan in the *glp-4* background and a statistically significant decrease in the *glp-4; daf-2* background.

Figure 7.2 *ins-7* and *ins-35* RNAi lifespans at 25°C using NGM-lactose plates

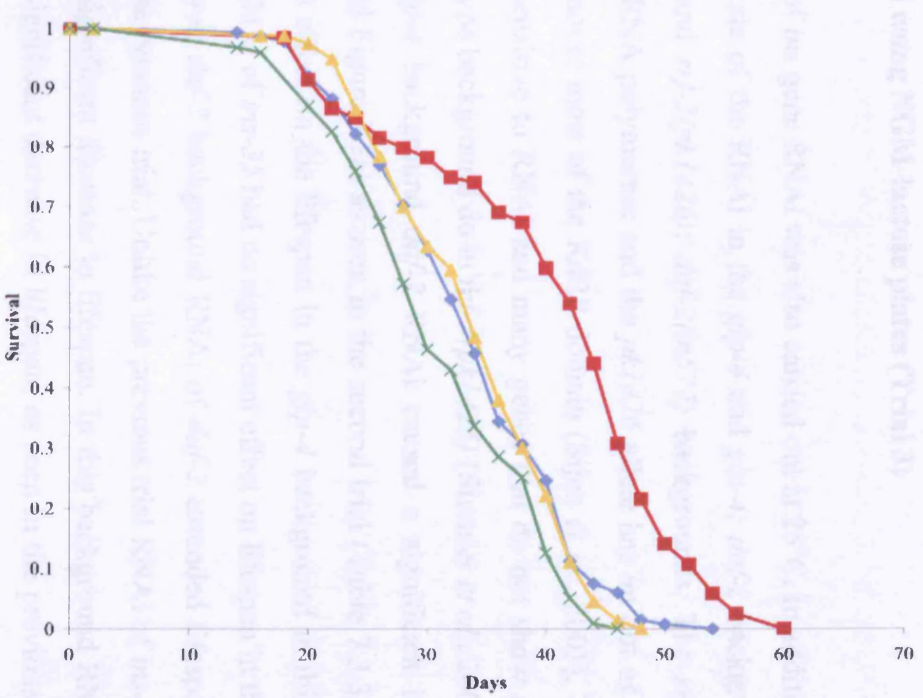
A Trial 2 *glp-4(bn2)* Lifespans at 25°C



RNAi of *ins-7* has no effect in *glp-4* background at 25°C. RNAi of *ins-35* leads to slight increase in mean lifespan in this background (p = 0.009). RNAi of *daf-2* causes a large increase in mean lifespan as expected.

L4440 Control
daf-2 RNAi
ins-7 RNAi
ins-35 RNAi

B Trial 2 *glp-4(bn2)*; *daf-2(m577)* Lifespans at 25°C



RNAi of *ins-7* has no effect in *glp-4*; *daf-2* background at 25°C. RNAi of *ins-35* leads to slight decrease in mean lifespan in this background (p = 0.005). RNAi of *daf-2* causes a large increase in mean lifespan as expected.

ins gene RNAi using NGM-lactose plates (Trial 3)

The third trial of *ins* gene RNAi was also carried out at 25°C. In addition to carrying out a second replicate of the RNAi in the *glp-4* and *glp-4; daf-2* backgrounds I also used *rrf-3(pk1426)* and *rrf-3(pk1426)*; *daf-2(m577)* backgrounds. The *rrf-3* gene encodes RNA-directed RNA polymerase and the *pk1426* allele has an out of frame deletion in exon 4 that removes most of the RdRP domain (Sijen *et al.*, 2001). The *rrf-3(pk1426)* allele is hypersensitive to RNAi and many genes that do not show a phenotype upon RNAi in wild type background do in *rrf-3(pk1426)* (Simmer *et al.*, 2002).

In the *glp-4* background *daf-2* RNAi caused a significant lifespan extension (Table 7.3.4 and Figure.7.3a) as seen in the second trial (Table 7.3.3). Likewise RNAi of *ins-7* had no effect on the lifespan in the *glp-4* background in this trial. Unlike the previous trial RNAi of *ins-35* had no significant effect on lifespan in this trial.

In the *glp-4*; *daf-2* background RNAi of *daf-2* extended lifespan, replicating the effect seen in the previous trial. Unlike the previous trial RNAi of *ins-7* in this trial lead to a small but significant decrease in lifespan. In this background RNAi of *ins-35* lead to a small but significant decrease in lifespan, as seen in the previous trial (Table 7.3.4 and Figure.7.3b).

In the hypersensitive strain *rrf-3* RNAi of *daf-2* also resulted in a large extension of lifespan compared to controls. RNAi of *ins-7* had no significant effect on lifespan in this background, whereas RNAi of *ins-35* lead to a slight decrease in the lifespan compared to controls (Table 7.3.4 and Figure.7.3c).

RNAi of *daf-2* also extended the lifespan of *rrf-3*; *daf-2* worms compared to controls. While RNAi of both *ins-7* and *ins-35* caused a small, but significant, decrease in the lifespan of worms in this background compare to controls (Table 7.3.4 and Figure.7.3d).

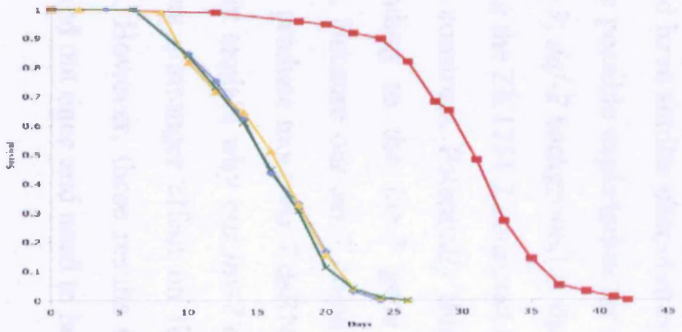
Table 7.3.4 *ins-7* and *ins-35* RNAi lifespans using NGM-lactose plates (Trial 3)

Genotype tested	RNAi construct	Mean Lifespan (Days)	Standard Error	*p*-value (log-rank)	N	Censor	% Control Lifespan
glp-4	L4440	16.35	0.43	N/A	97	3	100
glp-4	*daf-2*	30.97	0.53	<0.0001	100	0	189.41
glp-4	*ins-7*	16.41	0.45	0.8164	98	2	100.35
glp-4	*ins-35*	16.17	0.42	0.7244	99	1	98.92
glp-4; daf-2	L4440	34.23	0.92	N/A	94	6	100
glp-4; daf-2	*daf-2*	41.58	1.51	<0.0001	97	3	121.48
glp-4; daf-2	*ins-7*	31.77	0.80	0.0204	97	3	92.83
glp-4; daf-2	*ins-35*	32.07	0.78	0.0459	97	3	93.71
rrf-3	L4440	15.88	0.42	N/A	87	14	100
rrf-3	*daf-2*	24.03	0.81	<0.0001	82	4	151.27
rrf-3	*ins-7*	15.05	0.50	0.4374	77	4	94.75
rrf-3	*ins-35*	14.26	0.48	0.0233	73	8	89.74
rrf-3; daf-2	L4440	30.97	0.63	N/A	93	7	100
rrf-3; daf-2	*daf-2*	34.87	0.61	<0.0001	97	5	112.59
rrf-3; daf-2	*ins-7*	27.45	0.63	<0.0001	99	1	88.62
rrf-3; daf-2	*ins-35*	27.80	0.79	0.0016	74	6	89.76

Effects of *ins-7* and *ins-35* RNAi on lifespan in *glp-4(bn2)*, *glp-4(bn2)*; *daf-2(m577)*, *rrf-3* and *rrf-3*; *daf-2* backgrounds using NGM-lactose plates at 25°C. RNAi of *daf-2* was included as positive control for lifespan extension in all genotypes. RNAi of *ins-7* has no effect on lifespan in either *glp-4* or *rrf-3* backgrounds. However, *ins-7* RNAi lead to small, but significant, decreases in mean lifespan in both *glp-4*; *daf-2* and *rrf-3*; *daf-2*. RNAi of ins-35 had no effect in the glp-4 background, but lead to small, but significant, decreases in mean lifespan in *glp-4*; *daf-2*, *rrf-3* and *rrf-3*; *daf-2*.

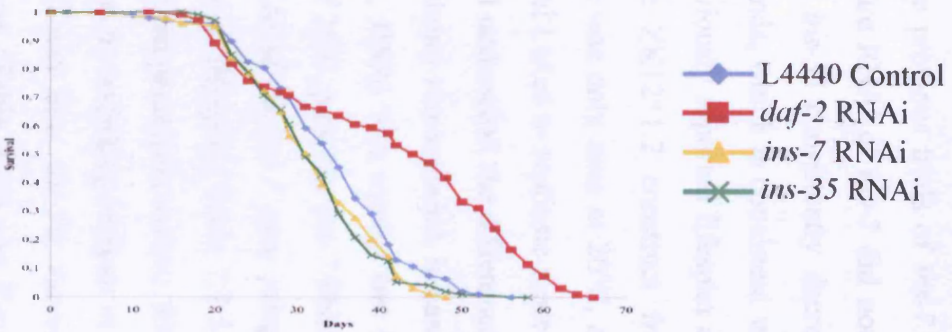
Figure 7.3 *ins-7* and *ins-35* RNAi lifespans at 25°C using NGM-lactose plates

A Trial 3 *glp-4(bn2)* Lifespans at 25°C



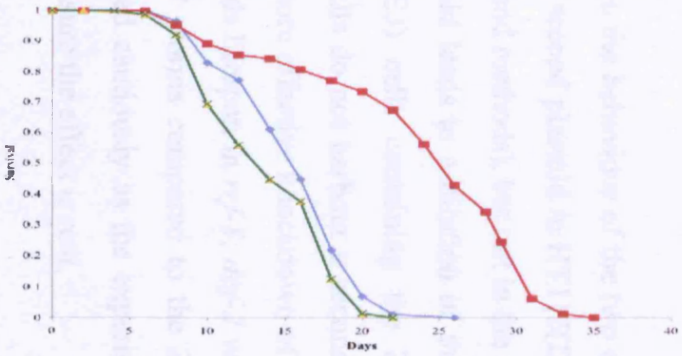
RNAi of *ins-7* and *ins-35* has no effect on mean lifespan in this trial. As expected RNAi of *daf-2* leads to lifespan extension.

B Trial 3 *glp-4(bn2)*; *daf-2(m577)* Lifespans at 25°C



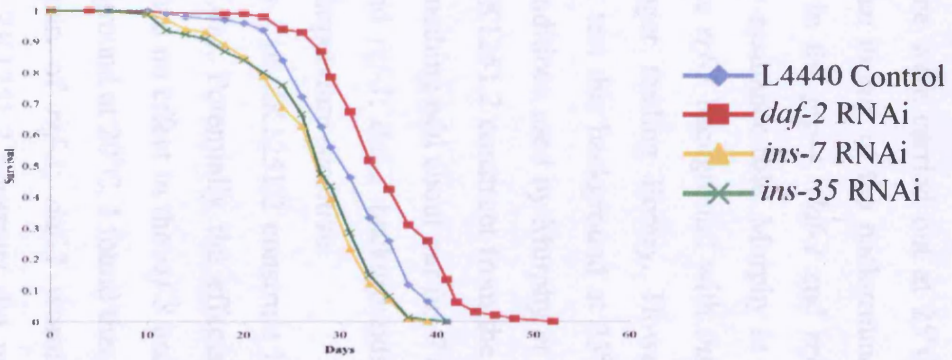
RNAi of *ins-7* and *ins-35* both lead to small decreases in mean lifespan (p = 0.02 and 0.05, respectively) in this background.

C Trial 3 *rrf-3(pk1426)* Lifespans at 25°C



RNAi of *ins-35* lead to a small decrease in mean lifespan (p = 0.02) in this background.

D Trial 3 *rrf-3(pk1426)*; *daf-2(m577)* Lifespans at 25°C



RNAi of *ins-7* and *ins-35* both lead to small decreases in mean lifespan (p = <0.0001 and 0.0016, respectively) in this background.

ins gene RNAi using IPTG plates (Trial 4)

All of the previous trials of *ins-7* and *ins-35* gene were carried out at 25°C. At this temperature RNAi of *ins-7* did not extend lifespan in any of the backgrounds tested. RNAi of *ins-35* consistently decreased lifespan in the *glp-4; daf-2* and *rrf-3*; *daf-2* backgrounds, which is consistent with a longevity assurance role. Murphy *et al* (2003) have previously reported lifespan extension in the *rrf-3* background with *ins-7* RNAi using the ZK1251.2 construct from the Ahringer feeding library. However, this extension was only seen at 20°C, as they did not test this background at 25°C. In the fourth trial I tried to replicate more closely the conditions used by Murphy *et al* (2003) to try and understand the difference. Firstly the ZK1251.2 construct from the Ahringer feeding library was included, in case there was something odd about our *ins-7* construct. Secondly, RNAi was carried out in the *rrf-3* and *rrf-3*; *daf-2* backgrounds at 20°C instead of 25°C; possibly *ins-7* RNAi effects are temperature sensitive.

RNAi of the *ins-7* gene using either the *ins-7* or ZK1251.2 construct lead to an extension of lifespan (Table 7.3.5 and Figure 7.4a). Potentially the effects of *ins-7* RNAi are temperature-sensitive. RNAi of *ins-35* had no effect in the *rrf-3* background. In addition to extending lifespan in the *rrf-3* background at 20°C, I found that our *ins-7* construct, and also *ins-35*, extended the lifespan of *rrf-3*; *daf-2* worms at this temperature (Table 7.3.6 and Figure 7.4b). The ZK1251.2 construct did not extend lifespan in this background. This is consistent with the findings of Murphy *et al* (2003). Although they used the *mu150* allele of *daf-2* whereas I used *m577*, both are class 1 alleles and have similar phenotypes.

One possible explanation for the difference in the behaviour of the two constructs in the *rrf-3*; *daf-2* background is the presence of a second plasmid in HT115(DE3) cells containing the ZK1251.2 construct (see Materials and methods), but not in the cells with the *ins-7* construct. Potentially this second plasmid leads to a dilution of the dsRNA corresponding to the *ins-7* gene in HT115(DE3) cells containing the ZK1251.2 construct. Because our *ins-7* construct bearing cells do not harbour a second plasmid they may produce more *ins-7* dsRNA leading to more effective knockdown of the gene. This might explain why our *ins-7* construct extends lifespan in *rrf-3*; *daf-2* worms and why it has a stronger effect on lifespan in *rrf-3* worms compared to the ZK1251.2 construct. However, these results should be treated cautiously as the experiment was only carried out once and need to be repeated to ensure the effect is real.

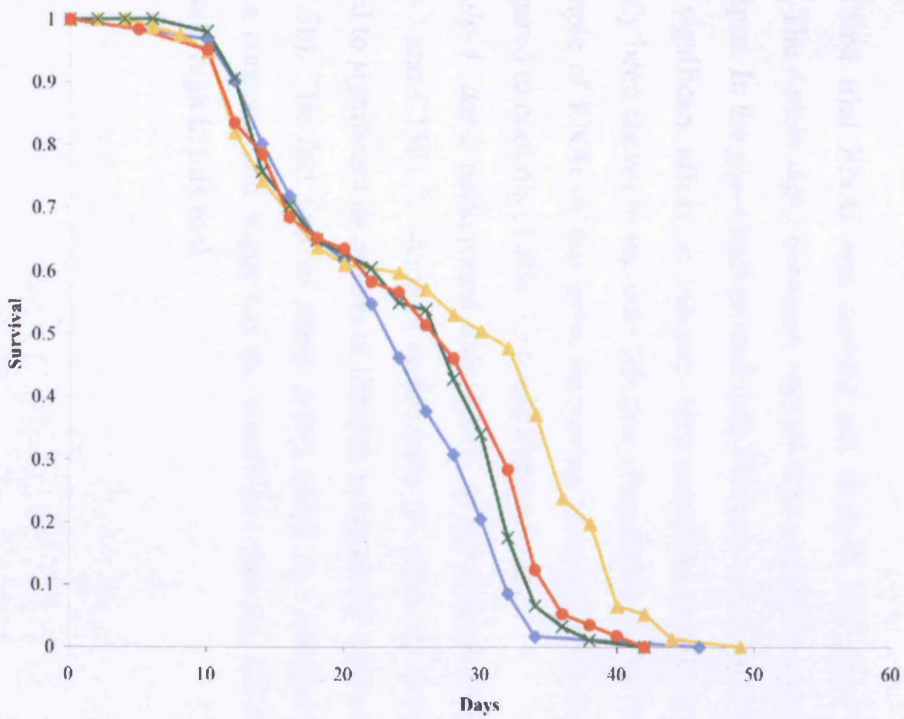
Table 7.3.5 *ins-7* and *ins-35* RNAi lifespans using IPTG plates (Trial 4)

Genotype tested	RNAi construct	Mean Lifespan (Days)	Standard Error	*p*-value (log-rank)	N	Censor	% Control Lifespan
rrf-3	L4440	23.46	1.03	N/A	59	1	100
rrf-3	ZK1251.2	25.16	1.26	0.0391	58	2	107.25
rrf-3	*ins-7*	27.27	1.35	0.0002	76	4	116.26
rrf-3	*ins-35*	24.71	0.90	0.1451	92	8	105.36
rrf-3; daf-2	L4440	26.86	0.58	N/A	90	10	100
rrf-3; daf-2	ZK1251.2	26.99	0.90	0.1489	65	15	100.48
rrf-3; daf-2	*ins-7*	32.24	0.73	<0.0001	91	9	120.04
rrf-3; daf-2	*ins-35*	29.59	0.60	0.0002	91	9	110.17

Effects of *ins-7* and *ins-35* RNAi on lifespan in *rrf-3* and *rrf-3*; *daf-2* backgrounds using IPTG plates at 20°C. Two different *ins-7* constructs were used in this trial, the *ins-7* construct used in the previous trials and the ZK1251.2 construct. ZK1251.2 is from the Ahringer feeding library and is the same construct that has been previously shown to extend lifespan by Murphy *et al.* (2003). In the *rrf-3* background both *ins-7* constructs extend mean lifespan by a small amount, with the *ins-7* construct having a greater effect than ZK1251.2. RNAi of *ins-35* has little effect on lifespan in the *rrf-3* background. In the *rrf-3*; *daf-2* background the *ins-7* construct, but not ZK1251.2, leads to mean lifespan extension, as does *ins-35* RNAi.

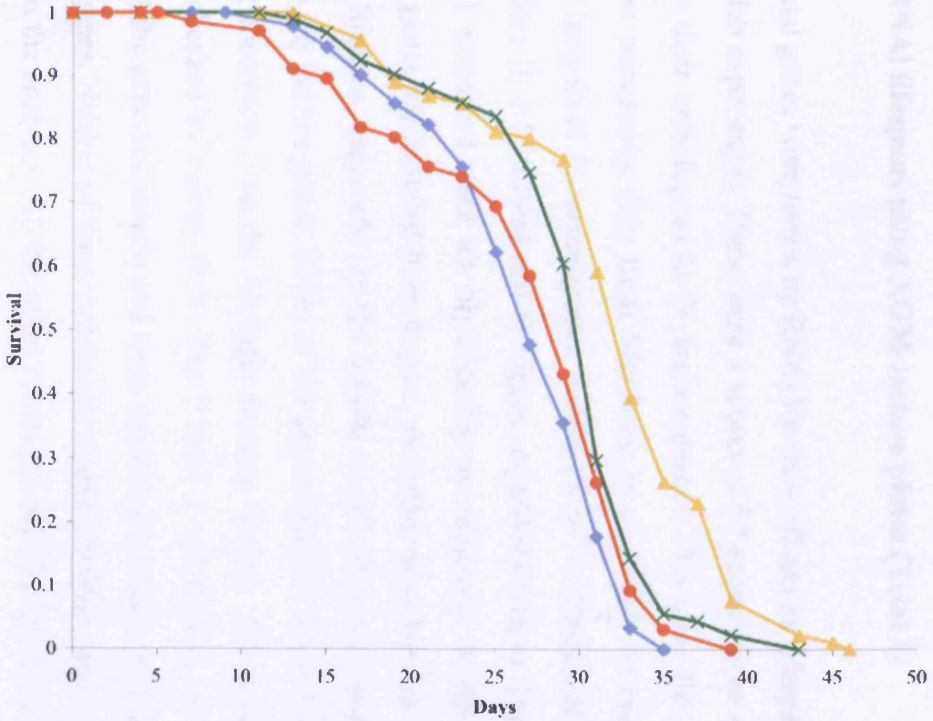
Figure 7.4 *ins-7* and *ins-35* RNAi lifespans at 20°C using NGM-lactose plates

A Trial 4 *rrf-3(pk1426)* Lifespans at 20°C



Both the ZK1251.2 and *ins-7* constructs lead to a small extension in mean lifespan (p = 0.04 and p = 0.0002, respectively), with *ins-7* having the greater effect. RNAi of *ins-35* has little effect.

B Trial 4 *rrf-3(pk1426)*; *daf-2(m577)* Lifespans at 20°C



L4440 Control
ins-7 RNAi
ins-35 RNAi
ZK1251.2 RNAi

The *ins-7* and *ins-35* constructs both lead to a small extension in mean lifespan (p = <0.0001 and p = 0.0002, respectively), with *ins-7* having the greater effect. ZK1251.2 has no effect in this background.

Additional RNAi lifespans using NGM-lactose plates (Trial 1)

Eight additional genes were tested by RNAi for their effects on lifespan by me as part of a communal lab experiment. These were a subset of 50 genes whose expression profiles are similar to their orthologues in *D. melanogaster*. This gene list was generated by comparing the microarray data from McElwee *et al* (2004) to microarray data that compared the long-lived *D. melanogaster* mutant $chico^{1}$/+ (Clancy *et al.*, 2001) to wild-type control flies (J. J. McElwee and D. Gems, unpublished data). The eight genes from this set that I examined were all significantly up regulated in *daf-2* mutants (Table 7.2.5). These genes are included here, in part, as further negative controls for the effects of RNAi on lifespan, especially as the L4440 vector does not produces dsRNA that correspond to any known genes. RNAi of the genes listed in Table 7.2.5 was carried out at 25°C using constructs from the Ahringer feeding library. All constructs had the size of the insert checked by colony PCR. The C36A4.2 and C55F2.1 PCR products were not the size of the expected inserts and were sequenced to see if they had any homology to the target genes. Neither of these constructs mapped to their target genes and so were excluded from the analyses. All the other constructs had the correct size insert (data not shown).

In this first trial RNAi was carried out in both the *glp-4* and *glp-4; daf-2* backgrounds. The Antebi *daf-2* construct was used as a positive control for the effect of RNAi on lifespan. In the *glp-4* background only RNAi of *daf-2* and H42K12.1, which is *pdk-1*, had a significant effect on lifespan. Hypomorphic mutations in the *pdk-1* gene have previously been shown to increase lifespan (Paradis *et al.*, 1999). However, this is the first example of RNAi of this gene increasing lifespan. Both caused extension of lifespan compared to controls (Table 7.3.6 and Figure 7.5a).

In the *glp-4; daf-2* background only RNAi of *daf-2* leads to lifespan extension. RNAi of *pdk-1* and C15F1.7, which is *sod-1*, have no effect on lifespan. All the other constructs lead to significant decreases in lifespan compared to the controls (Table 7.3.6 and Figure 7.5b). The fact that so many genes result in a similar small decrease in lifespan was a concern, and suggested the possibility that the L4440 control lifespan was artefactually high in this trial.

Table 7.3.6 Additional RNAi lifespans using NGM-lactose plates

Genotype tested	RNAi construct	Mean Lifespan (Days)	Standard Error	*p*-value (log-rank)	N	Censor	% Control Lifespan
glp-4	L4440	14.21	0.34	N/A	119	7	100
glp-4	*daf-2*	23.68	0.48	<0.0001	128	3	166.57
glp-4	T22B7.1 (*egl-13*)	14.48	0.34	0.6489	120	7	101.84
glp-4	C15F1.7 (*sod-1*)	14.44	0.46	0.7222	67	8	101.62
glp-4	H42K12.1 (*pdk-1*)	16.00	0.33	0.0001	150	5	112.55
glp-4	R11A5.4 (PEPCK)	13.43	0.31	0.1263	146	5	94.48
glp-4	C06A5.7 (*tmd-1*)	15.08	0.31	0.1306	128	2	106.08
glp-4	W01A11.1 (Epoxide hydroxylase)	14.68	0.36	0.2081	126	2	103.30
glp-4	K11E8.1 (*unc-43*)	14.64	0.31	0.3467	147	6	102.99
glp-4	F21A3.2 (Purple acid phosphatase)	14.18	0.31	0.8509	131	1	99.76
glp-4; daf-2	L4440	33.74	0.72	N/A	134	6	100
glp-4; daf-2	*daf-2*	40.69	1.00	<0.0001	121	5	120.59
glp-4; daf-2	T22B7.1 (*egl-13*)	25.62	0.52	<0.0001	121	3	75.94
glp-4; daf-2	C15F1.7 (*sod-1*)	33.05	0.68	0.3003	152	4	97.96
glp-4; daf-2	H42K12.1 (*pdk-1*)	32.98	0.67	0.2655	145	3	97.74
glp-4; daf-2	R11A5.4 (PEPCK)	25.70	0.59	<0.0001	119	7	76.17
glp-4; daf-2	C06A5.7 (*tmd-1*)	28.00	0.53	<0.0001	150	0	82.99
glp-4; daf-2	W01A11.1 (Epoxide hydroxylase)	27.64	0.57	<0.0001	118	7	81.92
glp-4; daf-2	K11E8.1 (*unc-43*)	28.05	0.50	<0.0001	139	6	83.14
glp-4; daf-2	F21A3.2 (Purple acid phosphatase)	28.30	0.48	<0.0001	142	11	83.87

Effects of RNAi of genes that are up-regulated in *glp-4*; *daf-2* double mutants compared to *glp-4 daf-16*; *daf-2* triple mutants using NGM-lactose plates at 25°C in *glp-4* and *glp-4*; *daf-2* backgrounds. In the *glp-4* background RNAi of H42K12.1 (*pdk-1*) leads to an extension of mean lifespan, all other genes, except *daf-2* control, have little effect. In the *glp-4*; *daf-2* background almost all the genes except H42K12.1 (*pdk-1*) and C15F1.7 (*sod-1*) lead to a decrease in mean lifespan, this suggests that the L4440 control may have been long-lived in this trial.

Additional RNAi lifespans using IPTG plates (Trial 2)

A second trial of the genes listed in Table 7.2.5 was carried out, this time using IPTG plates. RNAi was carried out at 25°C in an *rrf-3(pk1426)*; *daf-2(m577)* background. RNAi of R11A5.4 was not carried out in this trial due to an insufficient number of animals. Instead RNAi of R11A5.4 in both this background and *rrf-3(pk1426)* was carried in a separate third trial using IPTG plates (Table 7.3.7 and Figure 7.5c). In this trial only RNAi of *daf-2* lead to a significant increase in lifespan compared to control animals (Table 7.3.7 and Figure 7.5c). RNAi of *pdk-1* and *sod-1* lead to non-significant increases in lifespan. The only RNAi construct to reduce lifespan compared to controls was C06A5.7 (*tmd-1*), which while significant has a very small effect. The results of this trial support the view that the L4440 control in the first trial was long-lived.

Table 7.3.7 Additional RNAi lifespans using IPTG plates (Trial 2)

Genotype tested	RNAi construct	Mean Lifespan (Days)	Standard Error	*p*-value (log-rank)	N	Censor	% Control Lifespan
rrf-3; daf-2	L4440	29.27	0.98	N/A	98	2	100
rrf-3; daf-2	*daf-2*	32.62	0.92	0.0116	99	7	111.47
rrf-3; daf-2	T22B7.1 (*egl-13*)	30.10	0.90	0.7382	77	4	102.87
rrf-3; daf-2	C15F1.7 (*sod-1*)	31.65	0.85	0.279	95	6	108.16
rrf-3; daf-2	H42K12.1 (*pdk-1*)	31.14	0.72	0.9729	97	4	106.41
rrf-3; daf-2	C06A5.7 (*tmd-1*)	27.89	0.77	0.0027	77	4	95.29
rrf-3; daf-2	W01A11.1 (Epoxide Hydroxylase)	30.37	1.04	0.1945	100	0	103.78
rrf-3; daf-2	K11E8.1 (*unc-43*)	28.51	0.96	0.6218	96	1	97.43
rrf-3; daf-2	F21A3.2 (Purple acid phosphatase)	29.51	0.86	0.5868	100	2	100.84

Effects of RNAi of genes that are up-regulated in *glp-4*; *daf-2* double mutants compared to *glp-4 daf-16*; *daf-2* triple mutants using IPTG plates at 25°C in a *rrf-3*; *daf-2* background. This trial excludes R11A5.4. RNAi of C06A5.7 (*tmd-1*) led to a small suppression of mean lifespan. All other genes had little effect, except for the *daf-2* control.

RNAi of R11A5.4 (PEPCK) using IPTG at 25°C (Trial 3)

In this trial RNAi of R11A5.4 did not have a significant effect on lifespan in the *rrf-3* background compared to the L4440 control (Table 7.3.8 and Figure 7.6a). In the *rrf-3*; *daf-2* background RNAi of R11A5.4 leads to significant reduction in lifespan compared to the controls (Figure 7.6b). The verification of the effect of R11A5.4 RNAi on the Age phenotype of the *rrf-3*; *daf-2* mutant is interesting since it implies that increased gluconeogenesis contributes to the longevity of *daf-2* mutants.

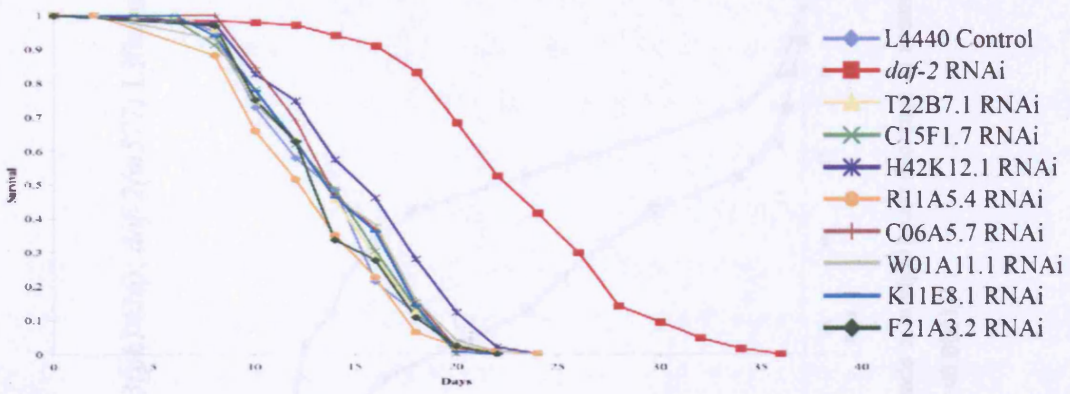
Table 7.3.8 RNAi of R11A5.4 (PEPCK) using IPTG at 25°C (Trial 3)

Genotype tested	RNAi construct	Mean Lifespan (Days)	Standard Error	*p*-value (log-rank)	N	Censor	% Control Lifespan
rrf-3	L4440	14.52	0.61	N/A	70	10	100
rrf-3	R11A5.4	13.41	0.48	0.1132	96	9	92.38
rrf-3; daf-2	L4440	33.23	0.72	N/A	96	4	100
rrf-3; daf-2	R11A5.4	27.19	0.71	<0.0001	101	4	81.82

Effect of R11A5.4 RNAi on lifespans in *rrf-3* and *rrf-3*; *daf-2* backgrounds at 25°C using IPTG plates. RNAi of R11A5.4 leads to a small non-significant decrease in lifespan in an *rrf-3* background and a significant decrease in mean lifespan in an *rrf-3*; *daf-2* background.

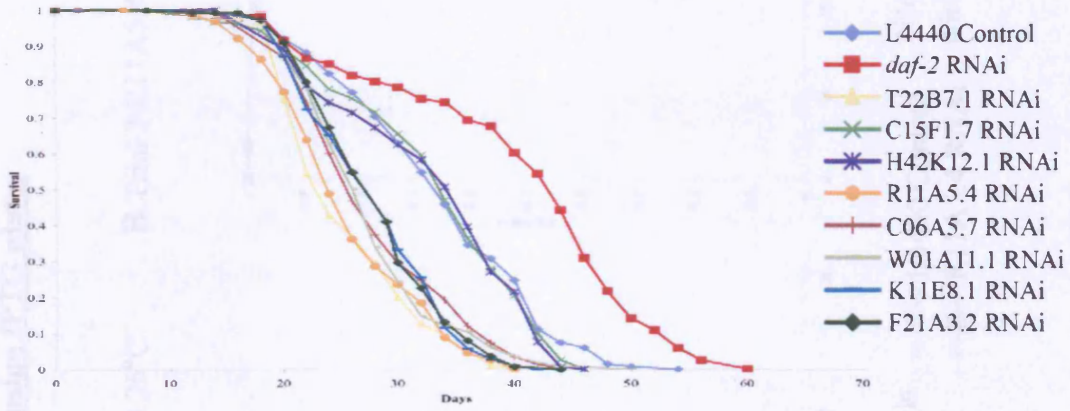
Figure 7.5 Additional RNAi lifespans at 25°C

A Trial 1 additional *glp-4(bn2)* Lifespans at 25°C on NGM-lactose plates



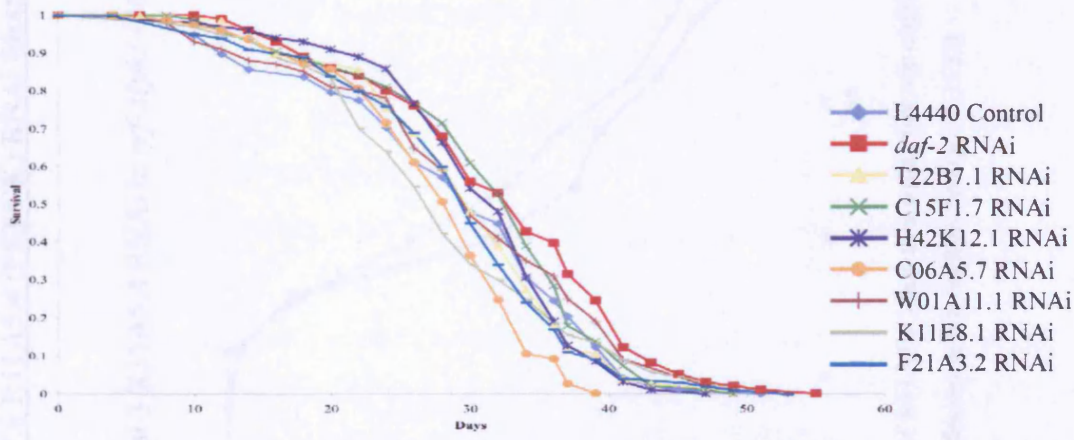
RNAi of H42K12.1 (*pdk-1*) leads to an increase in mean lifespan (p = <0.0001). No other treatment other than *daf-2* RNAi has a significant effect.

B Trial 1 additional *glp-4(bn2)*; *daf-2(m577)* Lifespans at 25°C on NGM-lactose plates



All RNAi treatments except H42K12.1 (*pdk-1*) and C15H1.7 (*sod-1*) lead to significant decreases in mean lifespan compared to control, suggesting that the L4440 control was actually long-lived.

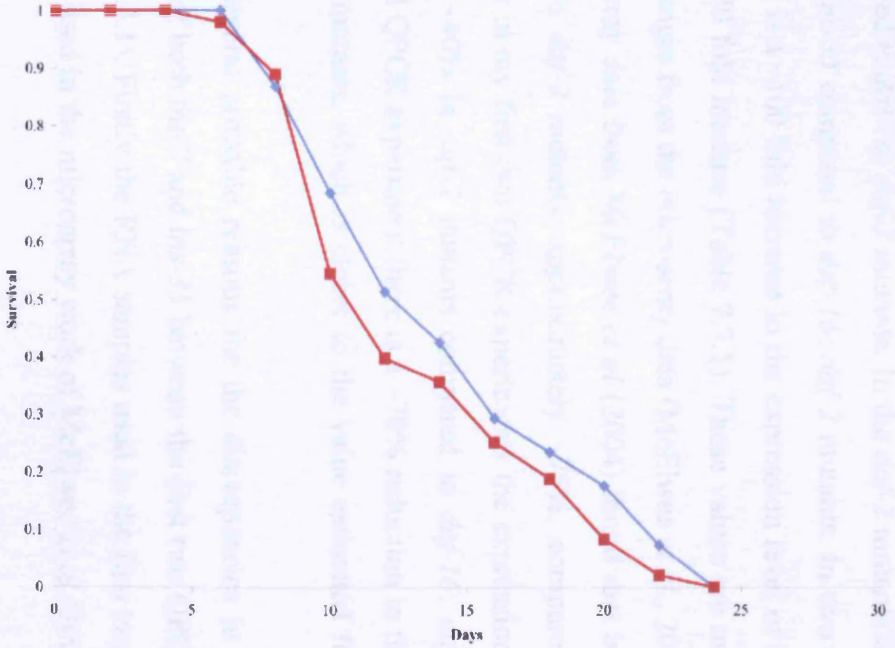
C Trial 2 additional *rrf-3(pk1426)*; *daf-2(m577)* Lifespans at 25°C on IPTG plates



In this trial only RNAi of C06A5.7 (*tmd-1*) had any effect on lifespan, other than the *daf-2* control. This treatment lead to small but significant decrease in lifespan (p = 0.003).

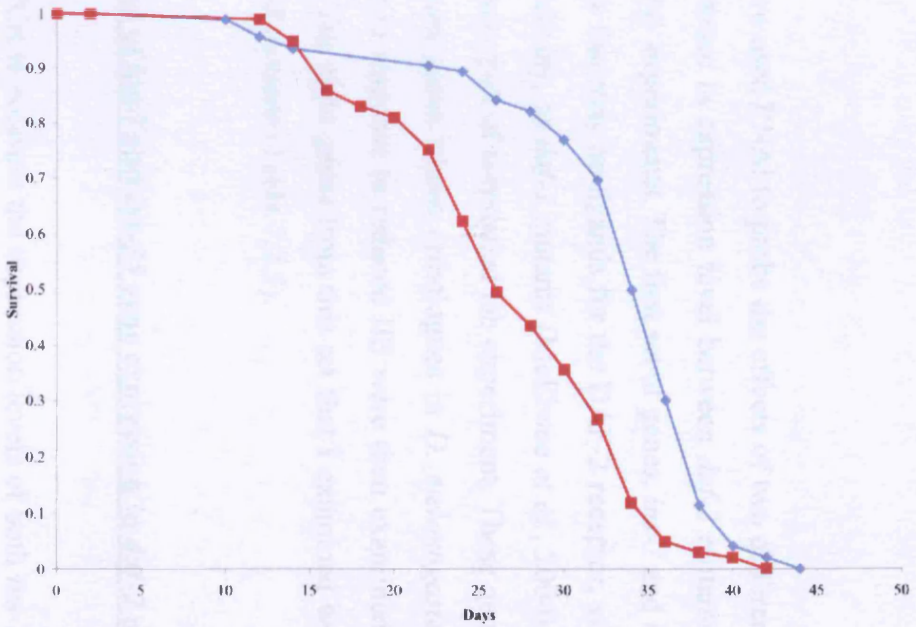
Figure 7.6 R11A5.4 (PEPCK) RNAi lifespans at 25°C using IPTG plates

A Trial 3 R11A5.4 RNAi in *rrf-3(pk1426)* Lifespans at 25°C



RNAi of R11A5.4 leads to a small, but non-significant, decrease in mean lifespan in this background ($p = 0.1132$)

B Trial 3 R11A5.4 RNAi in *rrf-3(pk1426)*; *daf-2(m577)* Lifespans at 25°C



L4440 Control
R11A5.4 RNAi

RNAi of R11A5.4 leads to a significant decrease in mean lifespan in this background ($p = <0.0001$)

7.4 Discussion

In this chapter I have used RNAi to probe the effects of two different sets of genes that show marked difference in expression level between *daf-2* mutants and *daf-16*; *daf-2* mutants in microarray experiments. The first set of genes, *ins-7* and *ins-35*, encodes two insulin-like peptides that may be ligands for the DAF-2 receptor, which are down- and up-regulated, respectively, in *daf-2* mutants (McElwee *et al.*, 2004). The second set of genes I examined were part of communal lab experiment. These genes were a subset of the top 50 *C. elegans* genes whose orthologues in *D. melanogaster* exhibited similar expression changes in response to reduced IIS were then examined for their effect on lifespan in worms. The eight genes from this set that I examined were all significantly up regulated in *daf-2* mutants (Table 7.2.5).

7.4.1 Quantification of *ins-7* and *ins-35* gene expression in *daf-2* mutants

I used real-time QPCR to measure the expression levels of both *ins-7* and *ins-35* in *daf-2* mutants compared to *daf-16*; *daf-2* mutants. In the *daf-2* mutants there is a significant up regulation of *ins-35* compared to *daf-16*; *daf-2* mutants. In two out of three QPCR experiments there is a ~100 fold increase in the expression level of *ins-35*, while in the third there is a ~70 fold increase (Table 7.3.1). These values are much larger than the estimated fold-changes from the microarray data (McElwee *et al.*, 2004) (Table 7.1.1).

The microarray data from McElwee *et al* (2004) found that *ins-7* is significantly down regulated in *daf-2* mutants, approximately ~75%, compared to *daf-16*; *daf-2* mutants. However in my first two QPCR experiments the expression levels of *ins-7* are only reduced by ~40% in *daf-2* mutants compared to *daf-16*; *daf-2* mutants (Table 7.3.1). In the third QPCR experiment there is a ~70% reduction in the expression levels of *ins-7* in *daf-2* mutants, which is closer to the value estimated from the microarray data.

There are several probable reasons for the discrepancies in the values for the expression levels of both *ins-7* and *ins-35* between the first two QPCR experiments and the third (Table 7.3.1). Firstly the RNA samples used in the first two experiments were the same as those used in the microarray work of McElwee *et al* (2004) and were a year

old. By contrast, the RNA samples used in the third experiment were prepared just prior to use. If there had been greater RNA degradation in the *daf-16*; *daf-2* samples used in the first two experiments compared to the *daf-2* samples then it would have appeared in the QPCR that these samples had less *ins-7* and *ins-35* transcript than they should. This in turn would make the levels of *ins-7* transcript appear similar or not significantly reduced and artificially boost the transcript levels of *ins-35* in the *daf-2* samples compared to *daf-16*; *daf-2*. Because the RNA used in the third QPCR experiment was prepared more recently the measured levels of *ins-7* and *ins-35* in both the *daf-2* and *daf-16*; *daf-2* samples may be closer to the biological levels. This may explain why the *ins-7* level in *daf-2* mutants in this experiment is closer to the estimated level from the microarray data.

A possible reason for the large value for the standard deviations may be the use of *ama-1* as the endogenous control gene. There are two requirements for a good endogenous control gene. Firstly the expression of the gene should remain unaffected between experimental treatments and secondly its expression level should be similar to the target genes (Bustin, 2000). The microarray data in McElwee et al (2004) shows that the gene expression of *ama-1* in *daf-2* mutants and *daf-16*; *daf-2* mutants is not significantly different. It has also previously been shown that *ama-1* expression is not affected by post-embryonic development (Johnstone and Barry, 1996; Larminie and Johnstone, 1996). This means *ama-1* fulfils the first requirement for a good control gene. As for the second requirement, the expression levels of *ama-1* are ~100 lower than the levels of both *ins-7* and *ins-35* (data not shown). This is problematic as small variations in the *ama-1* level of each sample can have a disproportionate effect during the normalisation procedure.

A further reason for the discrepancies in the values is the general nature of real-time QPCR itself. All QPCR experiments have the same basic assumptions that the efficiency of the reverse-transcription reaction and the subsequent PCR reactions will be equal in all the samples. However this unlikely to be true which is why samples are generally run in triplicates. For a detailed review of problems of reproducibility of QPCR see Bustin (2000).

7.4.2 Effect of RNAi of *ins-7* and *ins-35* on lifespan

RNAi of the *ins-7* gene has previously been shown to extend the lifespan of *rrf-3(pk1426)* worms at 20°C (Murphy *et al.*, 2003). This result suggests that the INS-7 protein may be an agonist of the DAF-2 receptor. This is also consistent with the microarray data from McElwee *et al* (2004) and my QPCR experiments which show that *ins-7* is down regulated in *daf-2* mutants compared to *daf-16*; *daf-2* mutants. In contrast the *ins-35* gene is significantly up regulated in *daf-2* mutants by ~70-100 fold compared to *daf-16*; *daf-2* mutants (Table 7.3.1).

This startling up-regulation of *ins-35* may be interpreted in two different ways. Firstly the up-regulation of *ins-35* may represent the *C. elegans* equivalent of the hyperinsulinaemia often associated with mutations in the *h*IR protein (see Taylor *et al.*, 1992 for a review of the clinical features of *h*IR mutations). Potentially the INS-35 protein may act as agonist of the DAF-2 receptor. When IIS pathway signalling is reduced by mutation of DAF-2, the subsequent increase in DAF-16 activity may cause an increase in the transcription of *ins-35*, in a futile attempt to reduce DAF-16 activity.

However, given that the INS-7 protein seems to be a DAF-2 agonist and *ins-7* is down-regulated in *daf-2* mutants it seems unlikely that INS-35 is also an agonist of DAF-2. This then raises the possibility that INS-35 is an antagonist of DAF-2 and that *ins-35* up-regulation in *daf-2* mutants plays a role in the increased longevity of *daf-2* worms. This would mean that the INS-35 protein might function as a longevity hormone in *C. elegans*. In order to test these hypotheses I used RNAi to reduce expression of *ins-35*, and also *ins-7* as a positive control, and test their effects on lifespan in variety of genetic backgrounds.

Table 7.4.1 Summary of effects of *ins-7* and *ins-35* on lifespan in all trials

Genotype	Trial	RNAi treatment	Temperature	+/- % of control lifespan
N2	1	*ins-7*	25°C	- 0.01
N2	1	*ins-35*	25°C	+ 0.06
daf-2	1	*ins-7*	25°C	+ 4.12
daf-2	1	*ins-35*	25°C	+ 14.36
glp-4	2	*ins-7*	25°C	- 0.12
glp-4	**2**	***ins-35***	**25°C**	**+ 8.71**
glp-4; daf-2	2	*ins-7*	25°C	+ 1.03
glp-4; daf-2	**2**	***ins-35***	**25°C**	**- 7.83**
glp-4	3	*ins-7*	25°C	+ 0.35
glp-4	3	*ins-35*	25°C	- 1.18
glp-4; daf-2	**3**	***ins-7***	**25°C**	**- 7.17**
glp-4; daf-2	**3**	***ins-35***	**25°C**	**- 6.39**
rrf-3	3	*ins-7*	25°C	- 5.25
rrf-3	**3**	***ins-35***	**25°C**	**- 10.26**
rrf-3; daf-2	**3**	***ins-7***	**25°C**	**- 11.38**
rrf-3; daf-2	**3**	***ins-35***	**25°C**	**- 10.24**
rrf-3	**4**	**ZK1251.2**	**20°C**	**+ 7.25**
rrf-3	**4**	***ins-7***	**20°C**	**+ 16.26**
rrf-3	4	*ins-35*	20°C	+ 5.36
rrf-3; daf-2	4	ZK1251.2	20°C	+ 0.48
rrf-3; daf-2	**4**	***ins-7***	**20°C**	**+ 20.04**
rrf-3; daf-2	**4**	***ins-35***	**20°C**	**+ 10.17**

Summary of the effects of RNAi of *ins-7* and *ins-35* in different genetic backgrounds and temperatures. The magnitude of the change compared to the control in each trial is listed in the final column. Genes listed in **bold** have a *p*-value <0.05

The results of all trials on the effects of *ins-7* and *ins-35* RNAi on lifespan are summarised in Table 7.4.1. In general the effect of *ins-35* RNAi on lifespan in the *glp-4*

and *rrf-3* backgrounds was variable at 25°C. By contrast, RNAi of *ins-35* in *glp-4*; *daf-2* and *rrf-3*; *daf-2* backgrounds consistently leads to a ~10% reduction in lifespan at this temperature. This result is consistent with concept of INS-35 being an antagonist of the DAF-2 receptor and functioning as a longevity hormone.

Surprisingly, even though *ins-35* levels are ~70-100 fold higher in *daf-2* mutants compared to *daf-16*; *daf-2* mutants, RNAi of *ins-35* only leads to ~10% reduction of lifespan. There are several possible reasons for this result. Firstly the QPCR experiment and the microarray data from McElwee *et al* (2004) do not contain a wild-type control. This means that we do not know the expression level of *ins-35* in wild type worms. If the expression level in wild-type worms is the same as that seen in the *daf-2* mutants then the QPCR and microarray data could be interpreted as showing significant down-regulation of *ins-35* in *daf-16*; *daf-2* mutants. If this were the case then INS-35 expression would not correlate with increased lifespan and therefore would not be expected to function as longevity hormone (or at least not a major one, given that RNAi does suppress lifespan a little).

A second reason that might explain why RNAi of *ins-35* has such a small effect on lifespan is functional redundancy. There are 39 insulin-like peptides encoded in the *C. elegans* genome, of which only 23 were represented on the Affymetrix array. One of the genes missing from the array is *ins-36*, whose encoded peptide is the most similar to INS-35. Potentially INS-36 can partially compensate for a loss or reduction of INS-35 activity. It would be interesting to know how *ins-36* is regulated by IIS.

A third reason for this effect may be related to the penetrance of the RNAi. It is has been shown that RNAi is not very effective in silencing genes that are expressed in the neuronal cells of *C. elegans* (Timmons *et al.*, 2001). Given that the insulin-like peptides of *C. elegans* are likely to function as neuromodulators/ neurotransmitters it is likely that *ins-35* knockdown in neurones will not be very effective even with the lactose-based induction protocol. If the neurones are the critical target tissue for INS-35 then the lack of major effect with *ins-35* RNAi is not that surprising. This could be resolved by using an *ins-35* deletion allele.

RNAi of the *ins-7* gene at 25°C, unexpectedly, lead to a reduction in lifespan in both *glp-4*; *daf-2* and *rrf-3*; *daf-2* backgrounds, in two out of three trials. RNAi of *ins-7* has previously been shown to extend lifespan in an *rrf-3* background by Murphy *et al* (2003), using the ZK1251.2 construct. However, they did not test for the effects of *ins-7* RNAi on lifespan in *rrf-3* at 25°C. Two possible reasons as to why my results at 25°C

do not replicate the results of Murphy *et al* (2003) are the use of a different *ins-7* RNAi construct, and the use of a different temperature. To distinguish between these two factors I repeated my RNAi experiments at 20°C, including the ZK1251.2 construct as positive control and using 20°C. I found that RNAi of *ins-7* with both the ZK1251.2 and our *ins-7* construct lead to a significant extension of lifespan in *rrf-3* worms. This implies that it is not difference in construct but the difference in temperature that accounts for the lack of replication at 25°C.

Why RNAi of the *ins-7* gene should extend lifespan at 20°C, yet lead to a reduction in lifespan at 25°C remains unclear, and is further complicated by the fact that *ins-35* RNAi also seems to extend lifespan at 20°C. It is difficult to explain why two INS proteins that have opposite regulation by DAF-16 should have the same contradictory phenotypes at two different temperatures. One possible interpretation of this data is that INS protein function is temperature-dependent rather than expression-level dependent. Thus at 20°C, potentially both peptides act as agonists of the DAF-2 receptor, but at 25°C both become antagonists.

7.4.3 Additional RNAi lifespan assays:

As part of a larger screen carried out in our lab I also performed RNAi on eight additional genes to probe their role in regulating lifespan extension in *daf-2* mutants. These genes show significant up-regulation in *daf-2* mutants and *chico*^{1}/+ heterozygotes (J. J. McElwee and D. Gems, unpublished data). The initial hypothesis was that these genes might be crucial to the lifespan extension seen in *daf-2* mutants compared to wild type worms. It was therefore predicted RNAi of these genes might suppress the lifespan extension seen in *daf-2* mutants.

The first trial was carried out at 25°C using NGM-lactose plates in *glp-4* and *glp-4*; *daf-2* backgrounds. Only H42K12.1 RNAi resulted in a significant effect on lifespan in the *glp-4* background. The H42K12.1 gene corresponds to the *pdk-1* gene and increased the lifespan of *glp-4* worms compared to controls (Table 7.3.6). The *pdk-1* gene is known to be part of the *daf-2* IIS pathway and hypomorphic alleles of this gene increase lifespan (Paradis *et al.*, 1999). This is the first example of lifespan extension by RNAi of this gene.

In subsequent trials of RNAi of these genes in the *rrf-3*; *daf-2* background using traditional IPTG plates only two genes had consistent, significant effects on lifespan (Table 7.3.7 and 7.3.8). One of these is C06A5.7, which corresponds to *tmd-1*, which encodes a tropomodulin homologue. Tropomodulins are pointed end actin capping proteins that regulate various actin networks. The other gene is R11A5.4, which encodes a member of the phosphoenolpyruvate carboxykinase (PEPCK) protein family. Both of these genes reduce lifespan compared to controls, with R11A5.4 RNAi having the strongest effect. RNAi of R11A5.4 was also carried out in the *rrf-3* background and found to have no significant effect on lifespan (Table 7.3.8). This suggests that R11A5.4 contributes to the extension of lifespan seen in *daf-2* mutants. PEPCK is rate-determining enzyme in gluconeogenesis. Thus, increased gluconeogenesis might contribute towards longevity assurance.

Chapter 8

Final conclusions

8.1 Summary of this thesis project

The aim of this thesis was to further our understanding of the *C. elegans daf-2* gene and the protein it encodes. In this thesis I have presented data on the evolution of *daf-2* gene within three *Caenorhabditis* species, as well as in the human parasitic nematode *B. malayi*, which suggests that the C-terminal extension in the *Caenorhabditis* species has evolved to promote rapid growth (Chapter 4). I have used an integrated analysis of sequence, mutant and structural data for the DAF-2, *h*IR and *h*IGF-1R proteins to generate the hypothesis that the molecular basis for the existence of two phenotypic classes of *daf-2* allele may be due to asymmetric alterations in the signalling outputs of DAF-2 (Chapter 5). In chapter 6, I performed some simple tests of my hypothesis and found that class 1 alleles could develop certain class 2 traits, but not others, which suggests my hypothesis may be too simplistic. In chapter 7, I probed the role of the *ins-7* and *ins-35* genes in determining longevity. These genes show opposite regulation by the transcription factor DAF-16, yet I found RNAi of both genes extended lifespan in *rrf-3* mutants at 20°C and suppressed lifespan-extension in *rrf-3*; *daf-2* double mutants. Below I reflect on the extent to which some of the questions I asked in chapter 2 have been answered and suggest future work to investigate the questions arising from this thesis.

8.2 Function of the N- and C-terminal extensions of DAF-2

The DAF-2 protein contains large N- and C-terminal extensions compared to *h*IR (see Figure 4.3) (Kimura *et al.*, 1997). At the beginning of this thesis project it was unknown whether these features were unique to DAF-2, or a common feature of DAF-2 proteins in other nematodes. The N-terminal extension contains no homology to any known proteins and is unusual in that it actually precedes the signal sequence (see Figure 4.4). The C-terminal extension is similar to that found in *D. melanogaster* (Fernandez *et al.*, 1995; Ruan *et al.*, 1995), and may subserve some of the functions of IRS-like proteins, even though the *C. elegans* genome contains genes encoding IRS, *ist-1* (Wolkow *et al.*, 2002) and Shc-like homologues, F54A5.2 and T27F7.2 (Luzi *et al.*, 2000).

The *Cb*-DAF-2 and *Cr*-DAF-2 proteins from the closely related nematodes *C. briggsae* and *C. remanei* have homologous large N- and C-terminal extensions. However, the *Bm*-DAF-2 protein from the more distantly related *B. malayi* does not possess an N-terminal extension and its C-terminal extension is much smaller than that of the *Caenorhabditis* species (see Figure 4.3). This suggests that the N-terminal extension is a feature that has evolved in the *Caenorhabditis* species and so presumably must have some function, even though no mutations have been identified in this domain (Chapter 5). Interestingly, the N-terminal extension in all three *Caenorhabditids* contains two putative endoproteolytic cleavage sites just N-terminal of the signal sequence (Figure 4.4) and so may be cleaved off prior to cleavage of the signal peptide. Potentially this region may function as an additional plasma membrane-targeting sequence, however, no other protein in *C. elegans* has a similar sequence at its N-terminus. Another possibility is that sequence may be involved in an interaction with chaperone proteins required for processing of DAF-2.

An ideal way to test for a function for the N-terminal extension would be to express transgenic *daf-2* constructs lacking portions of the extension in a *daf-2(0)/ qC1* strain. If the extension does have a membrane-targeting or chaperone-interaction function then these constructs would not rescue the Daf-c phenotype of this strain, if however, the extension had no function this strain would not segregate dauers. In order to differentiate between the membrane-targeting or chaperone-interaction functions one would need a good DAF-2 antibody. Using immunolocalization techniques it might be possible to visualise whether the N-terminal deletion product was being targeted to the plasma membrane or whether it was remaining cytosolic. Using co-immunoprecipitation it might be possible to see if wild-type DAF-2 pulled down chaperonins that the deletion product did not.

No mutations were found within the C-terminal extension during this project (Chapter 5). However, this thesis does provide the first direct evidence of a role in signal transduction for the C-terminal extension. This is shown in chapter 6, where in the search for a null allele of *daf-2* I found that the *tm1236* allele, which has an in-frame deletion within the TK domain and still possesses the C-terminal extension is fully suppressed by *daf-16(0)* at 25°C, whereas the *m65* nonsense allele, which is predicted to produce a receptor truncated within the TK domain, is 100% embryonic lethal at this temperature. This implies that DAF-2 has a DAF-16-independent role during embryogenesis and that the *tm1236* receptor has more activity than the *m65* receptor,

even though both proteins are predicted to kinase-dead, as *daf-16(mgDf50)*; *daf-2(tm1236)* animals develop into fertile adults.

In this project I also found that part of the C-terminal extensions in all three *Caenorhabditis* DAF-2 proteins is made of 3 repeats of a peptide sequence (Figure 4.6). Each of these repeats is encoded by a single exon, in all three species, implying that this feature has evolved by exon duplication. Moreover, these repeated exons appear to be related to the terminal exon of the *Bm-daf-2* gene, whose C-terminal extension is shorter than that of the *Caenorhabditids*. In chapter 4 I put forward the hypothesis that the extended C-terminal extension in the *Caenorhabditids* was an adaptive trait that had evolved to allow more rapid growth and development, consequently leading to shorter lifespan (Figure 4.7).

In order to test this theory a greater number of DAF-2 sequences, from a mixture of free-living and parasitic, as well as from long-lived and short-lived nematodes, need to be analysed to trace the evolution of the C-terminal extension. Transgenic *daf-2* constructs with targeted deletions of each of the repeats could be expressed in *daf-2(0)* strains. These would be expected to rescue the non-conditional Daf-c phenotype of the *daf-2(0)* allele and lead to an extension in lifespan compared to N2. A second possible way to test this theory would be to express these deletion constructs, along with wild-type DAF-2 and *Bm*-DAF-2, in a heterologous expression system. It might then be possible to assay the level of PIP3 production upon ligand binding by each receptor. By taking into account the expression levels of the receptors for each construct it should be possible to determine how much PIP3 is produced by each receptor. Wild-type DAF-2 would be expected to produce the most, followed by the deletion constructs, with *Bm*-DAF-2 expected to produce the least.

8.3 Molecular basis for the class 1 and 2 phenotypes

At the beginning of this project the simple working hypothesis for the molecular nature of the class 1/ class 2 distinction was that class 1 alleles had mutations in the ligand-binding domain, while class 2 alleles had mutations in the TK domain or hyper-conserved residues in the CR domain. This was based on the limited information available for several classified alleles sequenced by Kimura *et al.* (1997). However, sequencing of more alleles revealed a far more complex distribution of mutations across

the receptor, for both class 1 and 2 alleles, than previously assumed (see Chapter 5 and Figure 5.1). A comparative analysis of mutations in DAF-2 with those in *h*IR and *h*IGF-1R suggested the alternative hypothesis that class 1 alleles had mutations that affected levels of the receptor at the plasma membrane, but did not affect their function. While class 2 alleles had mutations that also affected their membrane levels, but crucially led to asymmetric signalling through the receptor, frequently affecting the Ras/ MAP kinase pathway more than PI 3-kinase signalling.

However, some simple tests of this hypothesis performed in chapter 6 suggest that it is too simplistic. RNAi of the *daf-2* gene in class 1 *daf-2(m41)*; *daf-12(m20)* double mutants at 25°C causes animals to arrest as sterile young adults, while controls develop into fertile adults. Class 2 *daf-2(e1370)*; *daf-12(m20)* double mutants arrest in similar manner at 22.5°C on standard NGM plates (Gems *et al.*, 1998; Larsen *et al.*, 1995). However, RNAi of *daf-2* in *m41* single mutants does not cause the development of any class 2 traits. I also found that the addition of *daf-28(sa191)* into a class 1 background led to the development of the early larval arrest phenotype seen in many class 2 alleles at 25°C, but crucially the class 1 *daf-2*; *daf-28* double mutants did not develop Eat and Unc phenotypes at this temperature, which are hallmark traits for class 2 alleles of *daf-2*. Potentially some phenotypes classically associated with class 2 alleles, such as the sterile adult arrest in a *daf-12* background and early larval arrest, may actually be associated with reduced levels of receptor and not asymmetric signalling. While other class 2 traits such as the development of the Eat and Unc phenotypes, which have never been seen in class 1 alleles, may be associated with asymmetric signalling. This is supported by the fact that activated Ras (LET-60 gain-of-function) signalling in a class 2 background suppresses the Eat phenotype (Nanji *et al.*, 2005).

I think that to really understand the molecular basis behind the class 1/ class 2 distinction we need to know which *daf-2* phenotypes are produced by reduction of receptors levels and which asymmetric signalling produces. Both the *daf-2* RNAi, in single and *daf-2*; *daf-12* double mutants, and the *daf-2*; *daf-28* experiments need to be repeated in a broader range of class 1 and 2 alleles to ensure that the effects seen in this thesis are not idiosyncratic to the particular alleles used in my experiments. Also class 1 alleles of *daf-2* should be tested for the development of Eat and Unc traits using RNAi of the components of Ras/ MAP kinase signalling.

8.4 Function of the insulin-like peptides in *C. elegans*

The *C. elegans* genome contains 39 genes that encode insulin-like peptides, which can be divided into four structural sub-types (Figure 1.8). One possible hypothesis explaining the existence of some putative ligands for the DAF-2 receptor is that some of these peptides behave as classical agonists, while others are actually antagonists (Pierce *et al.*, 2001). The *ins-7* and *ins-35* genes are oppositely regulated by the transcription factor DAF-16 in a *daf-2* mutant background(McElwee *et al.*, 2004). The *ins-7* gene is down-regulated by ~40% and *ins-35* is up-regulated 100-fold in *daf-2* mutants (Chapter 7). INS-7 is expected to behave as an agonist of DAF-2 and INS-35 an antagonist. This expectation is partially confirmed by the finding that *ins-7* RNAi extends lifespan in an *rrf-3(pk1426)* background at 20°C (Murphy *et al.*, 2003).

Based on this previous data I anticipated that INS-35 might function as longevity hormone and RNAi of the *ins-35* gene in a *daf-2(m577)* background might suppress its lifespan extension. However, RNAi of both genes seemed to extend lifespan at 20°C in a variety of genetic backgrounds and slightly suppress lifespan in both *glp-4*; *daf-2* and *rrf-3*; *daf-2* double mutants. This seemingly contradictory result may be due to three factors. The first may be the inherent variability of RNAi through feeding of *E. coli* expressing dsRNA of the intended target gene. The second factor is the scale of genetic redundancy given that there are 39 insulin-like peptide genes. The third factor is the temperature-sensitive nature of the dauer formation and longevity phenotypes in *C. elegans*. Potentially INS peptides can act as agonists at lower temperatures and antagonists at higher temperatures.

In order to further our understanding of the functions of *ins-7* and *ins-35*, it may be better to perform the RNAi experiments in a *daf-28(sa191)* background. This gene encodes mutant insulin-like peptide that is thought to act as a poisonous ligand for DAF-2 by forming complexes with other INS peptides and rendering them incapable of binding to the receptor (Li *et al.*, 2003). In addition, this allele has a weak lifespan extension of ~10-13%, therefore this background may be sensitized to the subtle effects of *ins-7* and *ins-35* RNAi. Recently the Japanese knockout consortium has produced two knockout alleles of *ins-7* (*tm1907* and *tm2001*), it would be beneficial to test both for lifespan extension compared to N2 at a range of temperatures. Other *ins* genes that knockouts are available for are *ins-1, -6, -11, -20, -23, -26* and *-29* of which *ins-11* and

ins-23 show significant alteration in expression levels between *daf-2* and *daf-16*; *daf-2* mutants (Table 7.1.1) (McElwee *et al.*, 2004). To fully elucidate the roles of all the INS peptides will require the construction strains carrying multiple deletions.

Bibliography

Abuzzahab, M. J., Schneider, A., Goddard, A., Grigorescu, F., Lautier, C., Keller, E., Kiess, W., Klammt, J., Kratzsch, J., Osgood, D., Pfaffle, R., Raile, K., Seidel, B., Smith, R. J. and Chernausek, S. D. (2003)
"IGF-I receptor mutations resulting in intrauterine and postnatal growth retardation" New England Journal of Medicine 349:2211-22

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990)
"Basic local alignment search tool" Journal of Molecular Biology 215:403-10

Antebi, A., Yeh, W. H., Tait, D., Hedgecock, E. M. and Riddle, D. L. (2000)
"*daf-12* encodes a nuclear receptor that regulates the dauer diapause and developmental age in *C. elegans*" Genes & Development 14:1512-27

Apfeld, J. and Kenyon, C. (1998)
"Cell nonautonomy of *C. elegans daf-2* function in the regulation of diapause and life span" Cell 95:199-210

Bajaj, M., Waterfield, M. D., Schlessinger, J., Taylor, W. R. and Blundell, T. (1987)
"On the tertiary structure of the extracellular domains of the epidermal growth factor and insulin receptors" Biochimica et biophysica acta 916:220-6

Bass, J., Chiu, G., Argon, Y. and Steiner, D. F. (1998)
"Folding of Insulin Receptor Monomers Is Facilitated by the Molecular Chaperones Calnexin and Calreticulin and Impaired by Rapid Dimerization" Journal of Cell Biology 141:637-646

Beanan, M. J. and Strome, S. (1992)
"Characterization of a germ-line proliferation mutation in *C. elegans*" Development 116:755-66

Bevan, A. P., Drake, P. G., Bergeron, J. J. M. and Posner, B. I. (1996)
"Intracellular signal transduction: The role of endosomes" Trends in Endocrinology and Metabolism 7:13

Bilan, P. J. and Yip, C. C. (1994)
"Unusual insulin binding to cells expressing an insulin receptor mutated at cysteine 524" Biochemical and Biophysical Research Communications 205:1891-8

Bird, D. M. and Riddle, D. L. (1989)
"Molecular cloning and sequencing of *ama-1*, the gene encoding the largest subunit of *Caenorhabditis elegans* RNA polymerase II" Molecular and Cellular Biology 9:4119-30

Blaxter, M. L., De Ley, P., Garey, J. R., Liu, L. X., Scheldeman, P., Vierstraete, A., Vanfleteren, J. R., Mackey, L. Y., Dorris, M., Frisse, L. M., Vida, J. T. and Thomas, W. K. (1998)
"A molecular evolutionary framework for the phylum Nematoda" Nature 392:71-5

Blom, N., Gammeltoft, S. and Brunak, S. (1999)
"Sequence and structure-based prediction of eukaryotic protein phosphorylation sites" Journal of Molecular Biology 294:1351-62

Bluher, M., Kahn, B. B. and Kahn, C. R. (2003)
"Extended longevity in mice lacking the insulin receptor in adipose tissue" Science 299:572-4

Blumenthal, T. and Steward, K. (1997)
"RNA Processing and Gene Structure" in *C. elegans* II Riddle, D. L., Blumenthal, T., Meyer, B. J. and Preiss, J. R. Cold Spring Harbor Laboratory Press

Bonafe, M., Barbieri, M., Marchegiani, F., Olivieri, F., Ragno, E., Giampieri, C., Mugianesi, E., Centurelli, M., Franceschi, C. and Paolisso, G. (2003)
"Polymorphic variants of insulin-like growth factor I (IGF-I) receptor and phosphoinositide 3-kinase genes affect IGF-I plasma levels and human longevity: cues for an evolutionarily conserved mechanism of life span control" Journal of Clinical Endocrinology and Metabolism 88:3299-304

Braeckman, B. P., Houthoofd, K. and Vanfleteren, J. R. (2001)
"Insulin-like signaling, metabolism, stress resistance and aging in *Caenorhabditis elegans*" Mechanisms of Ageing and Development122:673-93

Brenner, S. (1974)
"The Genetics of *Caenorhabditis elegans*" Genetics 77:71-94

Brogiolo, W., Stocker, H., Ikeya, T., Rintelen, F., Fernandez, R. and Hafen, E. (2001)
"An evolutionarily conserved function of the *Drosophila* insulin receptor and insulin-like peptides in growth control" Current Biology 11:213-21

Bustin, S. A. (2000)
"Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays" Journal of Molecular Endocrinology 25:169-193

Carrera, P., Cordera, R., Ferrari, M., Cremonesi, L., Taramelli, R., Andraghetti, G., Carducci, C., Dozio, N., Pozza, G., Taylor, S. I. and et al. (1993)
"Substitution of Leu for Pro-193 in the insulin receptor in a patient with a genetic form of severe insulin resistance" Human Molecular Genetics 2:1437-41

Cassada, R. C. and Russell, R. L. (1975)
"The Dauerlarva, a post-embryonic developmental variant of the nematode *Caenorhabditis elegans*" Developmental Biology 46:326-342

Ceresa, B. P., Kao, A. W., Santeler, S. R. and Pessin, J. E. (1998)
"Inhibition of Clathrin-Mediated Endocytosis Selectively Attenuates Specific Insulin Receptor Signal Transduction Pathways" Molecular and Cellular Biology 18:3862-3870

Cheatham, B. and Kahn, C. R. (1992)
"Cysteine 647 in the insulin receptor is required for normal covalent interaction between alpha- and beta-subunits and signal transduction" Journal of Biological Chemistry 267:7108-15

Cho, S., Jin, S. W., Cohen, A. and Ellis, R. E. (2004)
"A phylogeny of *Caenorhabditis* reveals frequent loss of introns during nematode evolution" Genome Researchearch 14:1207-20

Chow, J. C., Condorelli, G. and Smith, R. J. (1998)
"Insulin-like Growth Factor-I Receptor Internalization Regulates Signaling via the Shc/Mitogen-activated Protein Kinase Pathway, but Not the Insulin Receptor Substrate-1 Pathway" Journal of Biology Chemistry 273:4672-4680

Clancy, D. J., Gems, D., Harshman, L. G., Oldham, S., Stocker, H., Hafen, E., Leevers, S. J. and Partridge, L. (2001)
"Extension of life-span by loss of CHICO, a *Drosophila* insulin receptor substrate protein" Science 292:104-6

Coghlan, A. and Wolfe, K. H. (2004)
"Origins of recently gained introns in *Caenorhabditis*" Proceedings of the National Academy of Sciences USA 101:11362-7

C. elegans genome sequencing consortium. (1998)
"Genome sequence of the nematode *C. elegans*: a platform for investigating biology" Science 282:2012-2018

Cox, F. E. (2000)
"Elimination of lymphatic filariasis as a public health problem" Parasitology Today 16:135

Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M. and Barton, G. J. (1998)
"JPred: a consensus secondary structure prediction server" Bioinformatics 14:892-3

De Meyts, P. and Whittaker, J. (2002)
"Structural biology of insulin and IGF1 receptors: implications for drug design" Nature Reviews Drug Discovery 1:769-83

Dillin, A., Crawford, D. K. and Kenyon, C. (2002)
"Timing requirements for insulin/IGF-1 signaling in *C. elegans*" Science 298:830-834

Dlakic, M. (2002)
"A new family of putative insulin receptor-like proteins in *C. elegans*" Current Biology 12:R155-7

Duret, L., Guex, N., Peitsch, M. C. and Bairoch, A. (1998)
"New insulin-like proteins with atypical disulfide bond pattern characterized in *Caenorhabditis elegans* by comparative sequence analysis and homology modeling" Genome Researchearch 8:348-53

Dyrlov Bendtsen, J., Nielsen, H., von Heijne, G. and Brunak, S. (2004)
"Improved Prediction of Signal Peptides: SignalP 3.0" Journal of Molecular Biology 340:783

Ebina, Y., Edery, M., Ellis, L., Standring, D., Beaudoin, J., Roth, R. A. and Rutter, W. J. (1985a)
"Expression of a functional human insulin receptor from a cloned cDNA in Chinese hamster ovary cells" Proceedings of the National Academy of Sciences USA 82:8014-8

Ebina, Y., Ellis, L., Jarnagin, K., Edery, M., Graf, L., Clauser, E., Ou, J. H., Masiarz, F., Kan, Y. W. and Goldfine, I. D. (1985b)
"The human insulin receptor cDNA: the structural basis for hormone-activated transmembrane signalling" Cell 40:747-58

Eckel, J. and Reinauer, H. (1988)
"Involvement of hormone processing in insulin-activated glucose transport by isolated cardiac myocytes" Biochemical Journal 249:111-6

Ellis, L., Clauser, E., Morgan, D. O., Edery, M., Roth, R. A. and Rutter, W. J. (1986)
"Replacement of insulin receptor tyrosine residues 1162 and 1163 compromises insulin-stimulated kinase activity and uptake of 2-deoxyglucose" Cell 45:721

Falquet, L., Pagni, M., Bucher, P., Hulo, N., Sigrist, C. J., Hofmann, K. and Bairoch, A. (2002)
"The PROSITE database, its status in 2002" Nucleic Acids Research 30:235-8

Faure, R., Baquiran, G., Bergeron, J. J. and Posner, B. I. (1992)
"The dephosphorylation of insulin and epidermal growth factor receptors. Role of endosome-associated phosphotyrosine phosphatase(s)" Journal of Biological Chemistry 267:11215-21

Favelyukis, S., Till, J. H., Hubbard, S. R. and Miller, W. T. (2001)
"Structure and autoregulation of the insulin-like growth factor 1 receptor kinase" Nature Structural Biology 8:1058-63

Fernandez, R., Tabarini, D., Azpiazu, N., Frasch, M. and Schlessinger, J. (1995)
"The *Drosophila* insulin receptor homolog: a gene essential for embryonic development encodes two receptor isoforms with different signaling potential" EMBO journal 14:3373-84

Flores-Riveros, J. R., Sibley, E., Kastelic, T. and Lane, M. D. (1989)
"Substrate phosphorylation catalyzed by the insulin receptor tyrosine kinase. Kinetic correlation to autophosphorylation of specific sites in the beta subunit" Journal of Biological Chemistry 264:21557-72

Fodor, A., Riddle, D. L., Nelson, F. K. and Golden, J. W. (1983)
"Comparison of a new wild-type *Caenorhabditis briggsae* with laboratory strains of *C. briggsae* and *C. elegans*" Nematologica 29:203-217

Foster, J. M., Kumar, S., Ganatra, M. B., Kamal, I. H., Ware, J., Ingram, J., Pope-Chappell, J., Guiliano, D., Whitton, C., Daub, J., Blaxter, M. L. and Slatko, B. E. (2004) "Construction of bacterial artificial chromosome libraries from the parasitic nematode *Brugia malayi* and physical mapping of the genome of its *Wolbachia* endosymbiont" International Journal of Parasitology 34:733-46

Francis, R. and Waterston, R. H. (1991)
"Muscle cell attachment in *Caenorhabditis elegans*" Journal of Cell Biology 114:465-79

Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M. and Dubchak, I. (2004)
"VISTA: computational tools for comparative genomics" Nucleic Acids Research 32:W273-9

Fredriksson, R. and Schioth, H. B. (2005)
"The repertoire of G-protein-coupled receptors in fully sequenced genomes" Molecular Pharmacology 67:1414-25

Friedman, D. B. and Johnson, T. E. (1988a)
"A mutation in the *age-1* gene in *Caenorhabditis elegans* lengthens life and reduces hermaphrodite fertility" Genetics 118:75-86

Friedman, D. B. and Johnson, T. E. (1988b)
"Three mutants that extend both mean and maximum life span of the nematode, *Caenorhabditis elegans*, define the *age-1* gene" Journal of Gerontology 43:B102-9

Gardner, M. P., Gems, D. and Viney, M. E. (2004)
"Aging in a very short-lived nematode" Experimental Gerontology 39:1267-76

Garigan, D., Hsu, A. L., Fraser, A. G., Kamath, R. S., Ahringer, J. and Kenyon, C. (2002)
"Genetic analysis of tissue aging in *Caenorhabditis elegans*: a role for heat-shock factor and bacterial proliferation" Genetics 161:1101-12

Garrett, T. P., McKern, N. M., Lou, M., Frenkel, M. J., Bentley, J. D., Lovrecz, G. O., Elleman, T. C., Cosgrove, L. J. and Ward, C. W. (1998)
"Crystal structure of the first three domains of the type-1 insulin-like growth factor receptor" Nature 394:395-399

Garvey, W. T., Olefsky, J. M. and Marshall, S. (1986)
"Insulin induces progressive insulin resistance in cultured rat adipocytes. Sequential effects at receptor and multiple postreceptor sites" Diabetes 35:258-267

Gems, D. (2000)
"Longevity and ageing in parasitic and free-living nematodes" Biogerontology 1:289-307

Gems, D. and Riddle, D. L. (2000)
"Defining wild-type life span in *Caenorhabditis elegans*" Journal of Gerontology series A: Biological Sciences and Medical Sciences 55:B215-219

Gems, D., Sutton, A. J., Sundermeyer, M. L., Albert, P. S., King, K. V., Edgley, M. L., Larsen, P. L. and Riddle, D. L. (1998)
"Two pleiotropic classes of *daf-2* mutation affect larval arrest, adult behavior, reproduction and longevity in *Caenorhabditis elegans*" Genetics 150:129-55

Gerisch, B. and Antebi, A. (2004)
"Hormonal signals produced by DAF-9/cytochrome P450 regulate *C. elegans* dauer diapause in response to environmental cues" Development 131:1765-76

Gerisch, B., Weitzel, C., Kober-Eisermann, C., Rottiers, V. and Antebi, A. (2001)
"A hormonal signaling pathway influencing *C. elegans* metabolism, reproductive development, and life span" Developmental Cell 1:841-51

Ghedin, E., Wang, S., Foster, J. M. and Slatko, B. E. (2004)
"First sequenced genome of a parasitic nematode" Trends in Parasitology 20:151-3

Gil, E. B., Malone Link, E., Liu, L. X., Johnson, C. D. and Lees, J. A. (1999)
"Regulation of the insulin-like developmental pathway of *Caenorhabditis elegans* by a homolog of the PTEN tumor suppressor gene" Proceedings of the National Academy of Sciences USA 96:2925-30

Gilbert, W. (1978)
"Why genes in pieces?" Nature 271:501

Gish, W. and States, D. J. (1993)
"Identification of protein coding regions by database similarity search" Nature Genetics 3:266-72

Gregoire, F. M., Chomiki, N., Kachinskas, D. and Warden, C. H. (1998)
"Cloning and developmental regulation of a novel member of the insulin-like gene family in *Caenorhabditis elegans*" Biochemical and Biophysical Research Communations 249:385-90

GuhaThakurta, D., Schriefer, L. A., Waterston, R. H. and Stormo, G. D. (2004)
"Novel transcription regulatory elements in *Caenorhabditis elegans* muscle genes" Genome Research 14:2457-68

Hamer, I., Foti, M., Emkey, R., Cordier-Bussat, M., Philippe, J., De Meyts, P., Maeder, C., Kahn, C. R. and Carpentier, J. L. (2002)
"An arginine to cysteine(252) mutation in insulin receptors from a patient with severe insulin resistance inhibits receptor internalisation but preserves signalling events" Diabetologia 45:657-67

Hedges, S. B. (2002)
"The origin and evolution of model organisms" Nature Reviews Genetics3:838-49

Henderson, S. T. and Johnson, T. E. (2001)
"*daf-16* integrates developmental and environmental inputs to mediate aging in the nematode *Caenorhabditis elegans*" Current Biology 11:1975-80

Hertweck, M., Gobel, C. and Baumeister, R. (2004)
"*C. elegans* SGK-1 is the critical component in the Akt/PKB kinase complex to control stress response and life span" Developmental Cell 6:577-88

Hodgkin, J. (2005)
"Introduction to genetics and genomics" in WormBook, The *C. elegans* Research Community, WormBook

Hojlund, K., Hansen, T., Lajer, M., Henriksen, J. E., Levin, K., Lindholm, J., Pedersen, O. and Beck-Nielsen, H. (2004)
"A novel syndrome of autosomal-dominant hyperinsulinemic hypoglycemia linked to a mutation in the human insulin receptor gene" Diabetes 53:1592-8

Holzenberger, M., Dupont, J., Ducos, B., Leneuve, P., Geloen, A., Even, P. C., Cervera, P. and Le Bouc, Y. (2003)
"IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice" Nature 421:182-7

Hsin, H. and Kenyon, C. (1999)
"Signals from the reproductive system regulate the lifespan of *C. elegans*" Nature 399:362-6

Hsu, A. L., Murphy, C. T. and Kenyon, C. (2003)
"Regulation of aging and age-related disease by DAF-16 and heat-shock factor" Science 300:1142-5

Hubbard, S. R., Wei, L., Ellis, L. and Hendrickson, W. A. (1994)
"Crystal structure of the tyrosine kinase domain of the human insulin receptor" Nature 372:746-754

Imamura, T., Takata, Y., Sasaoka, T., Takada, Y., Morioka, H., Haruta, T., Sawa, T., Iwanishi, M., Hu, Y. G., Suzuki, Y. and et al. (1994)
"Two naturally occurring mutations in the kinase domain of insulin receptor accelerate degradation of the insulin receptor and impair the kinase activity" Journal of Biological Chemistry 269:31019-27

Imano, E., Kadowaki, H., Kadowaki, T., Iwama, N., Watarai, T., Kawamori, R., Kamada, T. and Taylor, S. I. (1991)
"Two patients with insulin resistance due to decreased levels of insulin-receptor mRNA" Diabetes 40:548-57

Jeong, P. Y., Jung, M., Yim, Y. H., Kim, H., Park, M., Hong, E., Lee, W., Kim, Y. H., Kim, K. and Paik, Y. K. (2005)
"Chemical structure and biological activity of the *Caenorhabditis elegans* dauer-inducing pheromone" Nature 433:541-5

Jia, K., Albert, P. S. and Riddle, D. L. (2002)
"DAF-9, a cytochrome P450 regulating *C. elegans* larval development and adult longevity" Development 129:221-31

Johnstone, I. L. (1999)
"Molecular Biology" in *C. elegans*: A Practical Approach Hope, I. A. Oxford University Press

Johnstone, I. L. and Barry, J. D. (1996)
"Temporal reiteration of a precise gene expression pattern during nematode development" Embo Journal 15:3633-9

Kadowaki, H., Kadowaki, T., Cama, A., Marcus-Samuels, B., Rovira, A., Bevins, C. L. and Taylor, S. I. (1990a)
"Mutagenesis of lysine 460 in the human insulin receptor. Effects upon receptor recycling and cooperative interactions among binding sites" Journal of Biological Chemistry 265:21285-96

Kadowaki, T., Kadowaki, H., Accili, D. and Taylor, S. I. (1990b)
"Substitution of lysine for asparagine at position 15 in the alpha-subunit of the human insulin receptor. A mutation that impairs transport of receptors to the cell surface and decreases the affinity of insulin binding" Journal of Biological Chemistry 265:19143-50

Kadowaki, T., Kadowaki, H., Accili, D., Yazaki, Y. and Taylor, S. I. (1991)
"Substitution of arginine for histidine at position 209 in the alpha-subunit of the human insulin receptor. A mutation that impairs receptor dimerization and transport of receptors to the cell surface" Journal of Biological Chemistry 266:21224-31

Kadowaki, T., Kadowaki, H., Rechler, M. M., Serrano-Rios, M., Roth, J., Gorden, P. and Taylor, S. I. (1990c)
"Five mutant alleles of the insulin receptor gene in patients with genetic forms of insulin resistance" Journal of Clinical Investigation 86:254-64

Kadowaki, T., Kadowaki, H. and Taylor, S. I. (1990d)
"A nonsense mutation causing decreased levels of insulin receptor mRNA: detection by a simplified technique for direct sequencing of genomic DNA amplified by the polymerase chain reaction" Proceedings of the National Academy of Sciences USA 87:658-62

Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D. P., Zipperlen, P. and Ahringer, J. (2003)
"Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi" Nature 421:231-7

Kamath, R. S., Martinez-Campos, M., Zipperlen, P., Fraser, A. G. and Ahringer, J. (2000)
"Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*" Genome Biology 2:1-10

Katic, M. and Kahn, C. R. (2005)
"The role of insulin and IGF-1 signaling in longevity" Cellular and Molecular Life Sciences (CMLS) 62:320-343

Kawano, T., Ito, Y., Ishiguro, M., Takuwa, K., Nakajima, T. and Kimura, Y. (2000)
"Molecular cloning and characterization of a new insulin/IGF-like peptide of the nematode *Caenorhabditis elegans*" Biochemical and Biophysical Research Communations 273:431-6

Kawano, T., Takuwa, K., Ishiguro, M., Nakajima, T. and Kimura, Y. (2003)
"Cloning and characterization of a *Caenorhabditis elegans* cDNA encoding a new insulin/IGF-like peptide" Bioscience, Biotechnology, and Biochemistry 67:2678-82

Kawashima, Y., Kanzaki, S., Yang, F., Kinoshita, T., Hanaki, K., Nagaishi, J., Ohtsuka, Y., Hisatome, I., Ninomoya, H., Nanba, E., Fukushima, T. and Takahashi, S. (2005)
"Mutation at cleavage site of insulin-like growth factor receptor in a short-stature child born with intrauterine growth retardation" Journal of Clinical Endocrinology and Metabolism 90:4679-87

Kenyon, C., Chang, J., Gensch, E., Rudner, A. and Tabtiang, R. (1993)
"A *C. elegans* mutant that lives twice as long as wild type" Nature 366:461-4

Khan, M. N., Baquiran, G., Brule, C., Burgess, J., Foster, B., Bergeron, J. J. and Posner, B. I. (1989)
"Internalization and activation of the rat liver insulin receptor kinase in vivo" Journal of Biological Chemistry 264:12931-12940

Khan, M. N., Savoie, S., Bergeron, J. J. and Posner, B. I. (1986)
"Characterization of rat liver endosomal fractions. In vivo activation of insulin-stimulable receptor kinase in these structures" Journal of Biological Chemistry 261:8462-8472

Kimura, K. D., Tissenbaum, H. A., Liu, Y. and Ruvkun, G. (1997)
"*daf-2*, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*" Science 277:942-6

Kiontke, K., Gavin, N. P., Raynes, Y., Roehrig, C., Piano, F. and Fitch, D. H. (2004)
"*Caenorhabditis* phylogeny predicts convergence of hermaphroditism and extensive intron loss" Proceedings of the National Academy of Sciences USA 101:9003-8

Klass, M. and Hirsh, D. (1976)
"Non-ageing developmental variant of *Caenorhabditis elegans*" Nature 260:523-5

Klein, H. H., Freidenberg, G. R., Matthaei, S. and Olefsky, J. M. (1987)
"Insulin receptor kinase following internalization in isolated rat adipocytes" Journal of Biological Chemistry 262:10557-10564

Knutson, V. P. (1991)
"Cellular trafficking and processing of the insulin receptor" FASEB Journal 5:2130-2138

Koppen, M., Simske, J. S., Sims, P. A., Firestein, B. L., Hall, D. H., Radice, A. D., Rongo, C. and Hardin, J. D. (2001)
"Cooperative regulation of AJM-1 controls junctional integrity in *Caenorhabditis elegans* epithelia" Nature Cell Biology 3:983-91

Krieger, M. J., Jahan, N., Riehle, M. A., Cao, C. and Brown, M. R. (2004)
"Molecular characterization of insulin-like peptide genes and their expression in the African malaria mosquito, *Anopheles gambiae*" Insect Molecular Biology 13:305-15

Krook, A., Humphreys, P. and O'Rahilly, S. (1997a)
"Molecular Mechanisms of Insulin Resistance" in Molecular Genetics of Endocrine Disorders, Thakker, R. V., Chapman and Hall Medical

Krook, A., Kumar, S., Laing, I., Boulton, A. J., Wass, J. A. and O'Rahilly, S. (1994)
"Molecular scanning of the insulin receptor gene in syndromes of insulin resistance" Diabetes 43:357-68

Krook, A., Moller, D. E., Dib, K. and O'Rahilly, S. (1996)
"Two naturally occurring mutant insulin receptors phosphorylate insulin receptor substrate-1 (IRS-1) but fail to mediate the biological effects of insulin. Evidence that IRS-1 phosphorylation is not sufficient for normal insulin action" Journal of Biological Chemistry 271:7134-40

Krook, A., Moller, D. E., Wass, J. A. and O'Rahilly, S. (1995)
"Functional studies of three naturally occurring insulin receptor mutations in the tyrosine kinase domain" Diabetes medecine 12:S21

Krook, A., Whitehead, J. P., Dobson, S. P., Griffiths, M. R., Ouwens, M., Baker, C., Hayward, A. C., Sen, S. K., Maassen, J. A., Siddle, K., Tavare, J. M. and O'Rahilly, S. (1997b)
"Two naturally occurring insulin receptor tyrosine kinase domain mutants provide evidence that phosphoinositide 3-kinase activation alone is not sufficient for the mediation of insulin's metabolic and mitogenic effects" Journal of Biological Chemistry 272:30208-14

Kublaoui, B., Lee, J. and Pilch, P. F. (1995)
"Dynamics of Signaling during Insulin-stimulated Endocytosis of Its Receptor in Adipocytes" Journal of Biological Chemistry 270:59-65

Labbe, J. C., Burgess, J., Rokeach, L. A. and Hekimi, S. (2000)
"ROP-1, an RNA quality-control pathway component, affects *Caenorhabditis elegans* dauer formation" Proceedings of the National Academy of Sciences USA 97:13233-8

Laemmli, U. K. (1970)
"Cleavage of structural proteins during the assembly of the head of bacteriophage T4" Nature 227:680-5

Lagueux, M., Lwoff, L., Meister, M., Goltzene, F. and Hoffmann, J. A. (1990)
"cDNAs from neurosecretory cells of brains of *Locusta migratoria* (Insecta, Orthoptera) encoding a novel member of the superfamily of insulins" European Journal of Biochemistry 187:249-254

Larminie, C. G. and Johnstone, I. L. (1996)
"Isolation and characterization of four developmentally regulated cathepsin B-like cysteine protease genes from the nematode *Caenorhabditis elegans*" DNA Cell Biology 15:75-82

Larsen, P. L., Albert, P. S. and Riddle, D. L. (1995)
"Genes that regulate both development and longevity in *Caenorhabditis elegans*" Genetics 139:1567-83

Le Roith, D. and Butler, A. A. (1999)
"Insulin-like Growth Factors in Pediatric Health and Disease" Journal of Clinical Endocrinology and Metabolism 84:4355-4361

Le Roith, D. and Zick, Y. (2001)
"Recent advances in our understanding of insulin action and insulin resistance" Diabetes Care 24:588-97

Li, W., Kennedy, S. G. and Ruvkun, G. (2003)
"*daf-28* encodes a *C. elegans* insulin superfamily member that is regulated by environmental cues and acts in the DAF-2 signaling pathway" Genes & Development 17:844-58

Lin, K., Dorman, J. B., Rodan, A. and Kenyon, C. (1997)
"*daf-16*: An HNF-3/forkhead family member that can function to double the life-span of *Caenorhabditis elegans*" Science 278:1319-22

Longo, N., Langley, S. D., Griffin, L. D. and Elsas, L. J., 2nd (1992)
"Reduced mRNA and a nonsense mutation in the insulin-receptor gene produce heritable severe insulin resistance" American Journal of Human Genetics 50:998-1007

Lu, K. and Guidotti, G. (1996)
"Identification of the cysteine residues involved in the class I disulfide bonds of the human insulin receptor: properties of insulin receptor monomers" Molecular Biology of the Cell 7:679-91

Luo, R. Z., Beniac, D. R., Fernandes, A., Yip, C. C. and Ottensmeyer, F. P. (1999)
"Quaternary structure of the insulin-insulin receptor complex" Science 285:1077-80

Luzi, L., Confalonieri, S., Di Fiore, P. P. and Pelicci, P. G. (2000)
"Evolution of Shc functions from nematode to human" Current Opinion in Genetics and Development 10:668-74

Maassen, J. A., Tobias, E. S., Kayserilli, H., Tukel, T., Yuksel-Apak, M., D'Haens, E., Kleijer, W. J., Fery, F. and van der Zon, G. C. (2003)
"Identification and functional assessment of novel and known insulin receptor mutations in five patients with syndromes of severe insulin resistance" Journal of Clinical Endocrinology and Metabolism 88:4251-7

Macaulay, S. L., Polites, M., Hewish, D. R. and Ward, C. W. (1994)
"Cysteine-524 is not the only residue involved in the formation of disulphide-bonded dimers of the insulin receptor" JBiochem Journal 303 (Pt 2):575-81

Maggi, D., Andraghetti, G., Carpentier, J. L. and Cordera, R. (1998)
"Cys860 in the extracellular domain of insulin receptor beta-subunit is critical for internalization and signal transduction" Endocrinology 139:496-504

Maggi, D., Barbetti, F. and Cordera, R. (1999)
"Role of proline 193 in the insulin receptor post-translational processing" Diabetologia 42:435-42

Maggi, D. and Cordera, R. (2001)
"Cys 786 and Cys 776 in the posttranslational processing of the insulin and IGF-I receptors" Biochemical and Biophysical Research Communations 280:836-41

Malone, E. A., Inoue, T. and Thomas, J. H. (1996)
"Genetic analysis of the roles of *daf-28* and *age-1* in regulating *Caenorhabditis elegans* dauer formation" Genetics 143:1193-205

Malone, E. A. and Thomas, J. H. (1994)
"A screen for nonconditional dauer-constitutive mutations in *Caenorhabditis elegans*" Genetics 136:879-86

Marino-Buslje, C., Mizuguchi, K., Siddle, K. and Blundell, T. L. (1998)
"A third fibronectin type III domain in the extracellular region of the insulin receptor family" FEBS Letters 441:331-6

Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F. and Sali, A. (2000)
"Comparative protein structure modeling of genes and genomes" Annual Reviews in Biophysical and Biomolecular Structures 29:291-325

McAlduff, D. E., Heng, Y. M. and Ottensmeyer, F. P. (2002)
"Freestanding lipid bilayers as substrates for electron cryomicroscopy of integral membrane proteins" Journal of Microscopy 205:113-7

McCulloch, D. and Gems, D. (2003)
"Evolution of male longevity bias in nematodes" Aging Cell 2:165-73

McElwee, J. J., Schuster, E., Blanc, E., Thomas, J. H. and Gems, D. (2004)
"Shared transcriptional signature in *Caenorhabditis elegans* Dauer larvae and long-lived *daf-2* mutants implicates detoxification system in longevity assurance" Journal of Biological Chemistry 279:44533-43

Mihaylova, V. T., Borland, C. Z., Manjarrez, L., Stern, M. J. and Sun, H. (1999)
"The PTEN tumor suppressor homolog in *Caenorhabditis elegans* regulates longevity and dauer formation in an insulin receptor-like signaling pathway" Proceedings of the National Academy of Sciences USA 96:7427-32

Mizuguchi, K., Deane, C. M., Blundell, T. L. and Overington, J. P. (1998)
"HOMSTRAD: a database of protein structure alignments for homologous families" Protein Science 7:2469-71

Moller, D. E., Benecke, H. and Flier, J. S. (1991)
"Biologic activities of naturally occurring human insulin receptor mutations. Evidence that metabolic effects of insulin can be mediated by a kinase-deficient insulin receptor mutant" Journal of Biological Chemistry 266:10995-1001

Moritz, W., Froesch, E. R. and Boni-Schnetzler, M. (1994)
"Functional properties of a heterozygous mutation (Arg1174-->Gln) in the tyrosine kinase domain of the insulin receptor from a type A insulin resistant patient" FEBS Letters 351:276-80

Morris, J. Z., Tissenbaum, H. A. and Ruvkun, G. (1996)
"A phosphatidylinositol-3-OH kinase family member regulating longevity and diapause in *Caenorhabditis elegans*" Nature 382:536-9

Mulhern, T. D., Booker, G. W. and Cosgrove, L. (1998)
"A third fibronectin-type-III domain in the insulin-family receptors" Trends in Biochemical Sciences 23:465-6

Murphy, C. T., McCarroll, S. A., Bargmann, C. I., Fraser, A., Kamath, R. S., Ahringer, J., Li, H. and Kenyon, C. (2003)
"Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*" Nature 424:277-83

Murray-Rust, J., McLeod, A. N., Blundell, T. L. and Wood, S. P. (1992)
"Structure and evolution of insulins: implications for receptor binding" Bioessays 14:325-31

Mynarcik, D. C., Williams, P. F., Schaffer, L., Yu, G. Q. and Whittaker, J. (1997)
"Identification of common ligand binding determinants of the insulin and insulin-like growth factor 1 receptors. Insights into mechanisms of ligand binding" Journal of Biological Chemistry 272:18650-5

Nakashima, N., Umeda, F., Yanase, T. and Nawata, H. (1995)
"Insulin resistance associated with substitution of histidine for arginine 252 in the alpha-subunit of the human insulin receptor: trial of insulin-like growth factor I injection therapy to enhance insulin sensitivity" Journal of Clinical Endocrinology and Metabolism 80:3662-7

Nanji, M., Hopper, N. A. and Gems, D. (2005)
"LET-60 RAS modulates effects of insulin/IGF-1 signaling on development and aging in *Caenorhabditis elegans*" Aging Cell 4:235-45

Ogg, S., Paradis, S., Gottlieb, S., Patterson, G. I., Lee, L., Tissenbaum, H. A. and Ruvkun, G. (1997)
"The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*" Nature 389:994-9

Ogg, S. and Ruvkun, G. (1998)
"The *C. elegans* PTEN homolog, DAF-18, acts in the insulin receptor-like metabolic signaling pathway" Molecular Cell 2:887-93

Ojamaa, K., Hedo, J. A., Roberts, C. T., Jr., Moncada, V. Y., Gorden, P., Ullrich, A. and Taylor, S. I. (1988)
"Defects in human insulin receptor gene expression" Molecular Endocrinology 2:242-7

Olson, T. S., Bamberger, M. J. and Lane, M. D. (1988)
"Post-translational changes in tertiary and quaternary structure of the insulin proreceptor. Correlation with acquisition of function" Journal of Biological Chemistry 263:7342-7351

Ottensmeyer, F. P., Beniac, D. R., Luo, R. Z. and Yip, C. C. (2000)
"Mechanism of transmembrane signaling: insulin binding and the insulin receptor" Biochemistry 39:12103-12

Paradis, S., Ailion, M., Toker, A., Thomas, J. H. and Ruvkun, G. (1999)
"A PDK1 homolog is necessary and sufficient to transduce AGE-1 PI3 kinase signals that regulate diapause in *Caenorhabditis elegans*" Genes & Development 13:1438-52

Paradis, S. and Ruvkun, G. (1998)
"*Caenorhabditis elegans* Akt/PKB transduces insulin receptor-like signals from AGE-1 PI3 kinase to the DAF-16 transcription factor" Genes & Development 12:2488-98

Petalcorin, M. I., Joshua, G. W., Agapow, P. M. and Dolphin, C. T. (2005)
"The fmo genes of *Caenorhabditis elegans* and *C. briggsae*: characterisation, gene expression and comparative genomic analysis" Gene 346:83-96

Pierce, S. B., Costa, M., Wisotzkey, R., Devadhar, S., Homburger, S. A., Buchman, A. R., Ferguson, K. C., Heller, J., Platt, D. M., Pasquinelli, A. A., Liu, L. X., Doberstein, S. K. and Ruvkun, G. (2001)
"Regulation of DAF-2 receptor signaling by human insulin and *ins-1*, a member of the unusually large and diverse *C. elegans* insulin gene family" Genes & Development 15:672-86

Pirola, L., Johnston, A. M. and Obberghen, E. (2004)
"Modulation of insulin action" Diabetologia 47:170-184

Podskalny, J., McElduff, A. and Gorden, P. (1984)
"Insulin receptors on Chinese hamster ovary (CHO) cells: altered insulin binding to glycosylation mutants" Biochemical and Biophysical Research Communations 125:70-5

Puig, O., Marr, M. T., Ruhf, M. L. and Tjian, R. (2003)
"Control of cell number by *Drosophila* FOXO: downstream and feedback regulation of the insulin receptor pathway" Genes & Development 17:2006-20

Riddle, D. L. and Albert, P. S. (1997)
"Genetic and Environmental Regulation of Dauer Larva Development" in *C. elegans* II Riddle, D. L., Blumenthal, T., Meyer, B. J. and Preiss, J. R. Cold Spring Harbor Laboratory press

Riddle, D. L., T., Meyer, B. J. and Preiss, J. R. (1997)
"*C. elegans* II" Cold Spring Harbor Laboratory press

Rosen, O. M., Herrera, R., Olowe, Y., Petruzzelli, L. M. and Cobb, M. H. (1983)
"Phosphorylation activates the insulin receptor tyrosine protein kinase" Proceedings of the National Academy of Sciences USA 80:3237-40

Rozen, S. and Skaletsky, H. (2000)
"Primer3 on the WWW for general users and for biologist programmers" Methods in Molecular Biology 132:365-86

Ruan, Y., Chen, C., Cao, Y. and Garofalo, R. S. (1995)
"The Drosophila insulin receptor contains a novel carboxyl-terminal extension likely to play an important role in signal transduction" Journal of Biological Chemistry 270:4236-43

Saltiel, A. R. and Kahn, C. R. (2001)
"Insulin signalling and the regulation of glucose and lipid metabolism" Nature 414:799-806

Sasaoka, T., Draznin, B., Leitner, J. W., Langlois, W. J. and Olefsky, J. M. (1994)
"Shc is the predominant signaling molecule coupling insulin receptors to activation of guanine nucleotide releasing factor and p21ras-GTP formation" Journal of Biological Chemistry 269:10734-8

Schaefer, E. M., Erickson, H. P., Federwisch, M., Wollmer, A. and Ellis, L. (1992)
"Structural organization of the human insulin receptor ectodomain" Journal of Biological Chemistry 267:23393-402

Schaffer, L. and Ljungqvist, L. (1992)
"Identification of a disulfide bridge connecting the alpha-subunits of the extracellular domain of the insulin receptor" Biochemical and Biophysical Research Communations 189:650-3

Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003)
"SWISS-MODEL: an automated protein homology-modeling server" Nucleic Acids Research 31:3381-3385

Scott, B. A., Avidan, M. S. and Crowder, C. M. (2002)
"Regulation of hypoxic death in *C. elegans* by the insulin/IGF receptor homolog DAF-2" Science 296:2388-91

Seino, S., Seino, M., Nishi, S. and Bell, G. I. (1989)
"Structure of the human insulin receptor gene and characterization of its promoter" Proceedings of the National Academy of Sciences USA 86:114-8

Shepherd, P. R., Nave, B. T. and Siddle, K. (1995)
"Insulin stimulation of glycogen synthesis and glycogen synthase activity is blocked by wortmannin and rapamycin in 3T3-L1 adipocytes: evidence for the involvement of phosphoinositide 3-kinase and p70 ribosomal protein-S6 kinase" Biochem Journal 305:25-28

Sherwood, O. D. (2004)
"Relaxin's Physiological Roles and Other Diverse Actions" Endocrine Reviews 25:205-234

Sijen, T., Fleenor, J., Simmer, F., Thijssen, K. L., Parrish, S., Timmons, L., Plasterk, R. H. and Fire, A. (2001)
"On the role of RNA amplification in dsRNA-triggered gene silencing" Cell 107:465-76

Simmer, F., Tijsterman, M., Parrish, S., Koushika, S. P., Nonet, M. L., Fire, A., Ahringer, J. and Plasterk, R. H. (2002)
"Loss of the putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi" Current Biology 12:1317-9

Skorokhod, A., Gamulin, V., Gundacker, D., Kavsan, V., Muller, I. M. and Muller, W. E. (1999)
"Origin of insulin receptor-like tyrosine kinases in marine sponges" Biology Bulletin 197:198-206

Smit, A. B., van Kesteren, R. E., Li, K. W., Van Minnen, J., Spijker, S., Van Heerikhuizen, H. and Geraerts, W. P. (1998)
"Towards understanding the role of insulin in the brain: lessons from insulin-related signaling systems in the invertebrate brain" Progress in Neurobiology 54:35-54

Sparrow, L. G., McKern, N. M., Gorman, J. J., Strike, P. M., Robinson, C. P., Bentley, J. D. and Ward, C. W. (1997)
"The disulfide bonds in the C-terminal domains of the human insulin receptor ectodomain" Journal of Biological Chemistry 272:29460-7

Stein, L. D., Bao, Z., Blasiar, D., Blumenthal, T., Brent, M. R., Chen, N., Chinwalla, A., Clarke, L., Clee, C., Coghlan, A., Coulson, A., D'Eustachio, P., Fitch, D. H., Fulton, L. A., Fulton, R. E., Griffiths-Jones, S., Harris, T. W., Hillier, L. W., Kamath, R., Kuwabara, P. E., Mardis, E. R., Marra, M. A., Miner, T. L., Minx, P., Mullikin, J. C., Plumb, R. W., Rogers, J., Schein, J. E., Sohrmann, M., Spieth, J., Stajich, J. E., Wei, C., Willey, D., Wilson, R. K., Durbin, R. and Waterston, R. H. (2003)
"The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics" PLoS Biology 1:E45

Strack, V., Stoyanov, B., Bossenmaier, B., Mosthaf, L., Kellerer, M. and Haring, H. U. (1997)
"Impact of mutations at different serine residues on the tyrosine kinase activity of the insulin receptor" Biochemical and Biophysical Research Communations 239:235-9

Sulston, J. E. and Horvitz, H. R. (1977)
"Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*" Developmental Biology 56:110-56

Sun, X. J., Rothenberg, P., Kahn, C. R., Backer, J. M., Araki, E., Wilden, P. A., Cahill, D. A., Goldstein, B. J. and White, M. F. (1991)
"Structure of the insulin receptor substrate IRS-1 defines a unique signal transduction protein" Nature 352:73-7

Suzuki, Y., Hatanaka, Y., Taira, M., Shimada, F., Hashimoto, N., Takayanagi, M., Taylor, S. I., Makino, H. and Yoshida, S. (1995)
"Insulin resistance associated with decreased levels of insulin-receptor messenger ribonucleic acid: evidence of a de novo mutation in the maternal allele" Journal of Clinical Endocrinology and Metabolism 80:1214-20

Takata, Y., Imamura, T., Haruta, T., Egawa, K., Takada, Y., Sawa, T., Yang, G. H. and Kobayashi, M. (1994)
"Leu 193 mutation in the cysteine rich region of the insulin receptor inhibits the cleavage of the insulin receptor precursor but not insulin binding" Biochemical and Biophysical Research Communations 203:763-7

Tatar, M., Kopelman, A., Epstein, D., Tu, M. P., Yin, C. M. and Garofalo, R. S. (2001)
"A mutant *Drosophila* insulin receptor homolog that extends life-span and impairs neuroendocrine function" Science 292:107-10

Tavare, J. M. and Siddle, K. (1993)
"Mutational analysis of insulin receptor function: consensus and controversy" Biochimica et biophysica acta 1178:21-39

Taylor, S. I., Cama, A., Accili, D., Barbetti, F., Quon, M. J., de la Luz Sierra, M., Suzuki, Y., Koller, E., Levy-Toledano, R., Wertheimer, E. and et al. (1992)
"Mutations in the insulin receptor gene" Endocrine Reviews 13:566-95

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994)
"CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice" Nucleic Acids Research 22:4673-80

Timmons, L., Court, D. L. and Fire, A. (2001)
"Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans" Gene 263:103-112

Timmons, L. and Fire, A. (1998)
"Specific interference by ingested dsRNA" Nature 395:854

Ullrich, A., Bell, J. R., Chen, E. Y., Herrera, R., Petruzzelli, L. M., Dull, T. J., Gray, A., Coussens, L., Liao, Y. C., Tsubokawa, M. and et al. (1985)
"Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes" Nature 313:756-61

van der Vorm, E. R., van der Zon, G. C., Moller, W., Krans, H. M., Lindhout, D. and Maassen, J. A. (1992)
"An Arg for Gly substitution at position 31 in the insulin receptor, linked to insulin resistance, inhibits receptor processing and transport" Journal of Biological Chemistry 267:66-71

van Heemst, D., Beekman, M., Mooijaart, S. P., Heijmans, B. T., Brandt, B. W., Zwaan, B. J., Slagboom, P. E. and Westendorp, R. G. J. (2005)
"Reduced insulin/IGF-1 signalling and human longevity" Aging Cell 4:79-85

Vanfleteren, J. R., Van de Peer, Y., Blaxter, M. L., Tweedie, S. A., Trotman, C., Lu, L., Van Hauwaert, M. L. and Moens, L. (1994)
"Molecular genealogy of some nematode taxa as based on cytochrome c and globin amino acid sequences" Molecular Phylogenetics and Evolution 3:92-101

Vowels, J. J. and Thomas, J. H. (1992)
"Genetic analysis of chemosensory control of dauer formation in *Caenorhabditis elegans*" Genetics 130:105-23

Wang, P. Y., Zhen, T. M., Wang, Z. Z., Gu, Z. F., Ren, S. P., Liu, L. H., Hou, L. W. and Liu, J. L. (1994)
"A ten-year observation on experimental infection of periodic *Brugia malayi* in man" Journal of Tropical Medicine and Hygiene 97:269-76

Ward, C. W. (1999)
"Members of the insulin receptor family contain three fibronectin type III domains" Growth Factors 16:315-22

Ward, C. W. and Garrett, T. P. (2001)
"The relationship between the L1 and L2 domains of the insulin and epidermal growth factor receptors and leucine-rich repeat modules" BMC Bioinformatics 2:4

Ward, C. W., Hoyne, P. A. and Flegg, R. H. (1995)
"Insulin and epidermal growth factor receptors contain the cysteine repeat motif found in the tumor necrosis factor receptor" Proteins 22:141-53

Wertheimer, E., Barbetti, F., Muggeo, M., Roth, J. and Taylor, S. I. (1994)
"Two mutations in a conserved structural motif in the insulin receptor inhibit normal folding and intracellular transport of the receptor" Journal of Biological Chemistry 269:7587-92

White, M. F., Maron, R. and Kahn, C. R. (1985)
"Insulin rapidly stimulates tyrosine phosphorylation of a Mr-185,000 protein in intact cells" Nature 318:183-6

White, M. F., Shoelson, S. E., Keutmann, H. and Kahn, C. R. (1988)
"A cascade of tyrosine autophosphorylation in the beta-subunit activates the phosphotransferase of the insulin receptor" Journal of Biological Chemistry 263:2969-80

Whitman, M., Downes, C. P., Keeler, M., Keller, T. and Cantley, L. (1988)
"Type I phosphatidylinositol kinase makes a novel inositol phospholipid, phosphatidylinositol-3-phosphate" Nature 332:644

Whittaker, J., Groth, A. V., Mynarcik, D. C., Pluzek, L., Gadsboll, V. L. and Whittaker, L. J. (2001)
"Alanine scanning mutagenesis of a type 1 insulin-like growth factor receptor ligand binding site" Journal of Biological Chemistry 276:43980-6

Wilkinson, T. N., Speed, T. P., Tregear, G. W. and Bathgate, R. A. (2005)
"Evolution of the relaxin-like peptide family" BMC Evolutionary Biology 5:14

Williams, P. F., Mynarcik, D. C., Yu, G. Q. and Whittaker, J. (1995)
"Mapping of an NH2-terminal ligand binding site of the insulin receptor by alanine scanning mutagenesis" Journal of Biological Chemistry 270:3012-6

Wolkow, C. A., Kimura, K. D., Lee, M. S. and Ruvkun, G. (2000)
"Regulation of *C. elegans* life-span by insulin-like signaling in the nervous system" Science 290:147-50

Wolkow, C. A., Munoz, M. J., Riddle, D. L. and Ruvkun, G. (2002)
"Insulin receptor substrate and p55 orthologous adaptor proteins function in the *Caenorhabditis elegans daf-2*/insulin-like signaling pathway" Journal of Biological Chemistry 277:49591-7

Wood, W.B (1988)
"The Nematode *Caenorhabditis elegans*" Cold Harbor Spring Laboratory

Wu, J. J. and Guidotti, G. (2002)
"Construction and characterization of a monomeric insulin receptor" Journal of Biological Chemistry 277:27809-17

Wu, J. J. and Guidotti, G. (2004)
"Proreceptor dimerization is required for insulin receptor post-translational processing" Journal of Biological Chemistry 279:25765-73

Yam, A., Hyun, T. and Li, W. (2001)
"Characterization of insulin-like growth factor I (IGF-I) receptor mutants for their effects on IGF-I- and interleukin 4-mediated DNA synthesis of 32D cells" Journal of Biological Chemistry 276:24409-13

Yokota, A., Moller, D. E. and Flier, J. S. (1990)
"Homozygous mutation at position 485 of the insulin receptor α -subunit in patient with lipodystrophy and severe insulin resistance." Diabetes 39:235A

Yoshida, I., Moto, K., Sakurai, S. and Iwami, M. (1998)
"A novel member of the bombyxin gene family: structure and expression of bombyxin G1 gene, an insulin-related peptide gene of the silkmoth Bombyx mori" Developmental Genes and Evolution 208:407-10

Yu, H. and Larsen, P. L. (2001)
"DAF-16-dependent and independent expression targets of DAF-2 insulin receptor-like pathway in Caenorhabditis elegans include FKBPs" Journal of Molecular Biology 314:1017-28